Science
AAAS

Supplementary Materials for

Diverse variola virus (smallpox) strains were widespread in northern Europe in the Viking Age

Barbara Mühlemann, Lasse Vinner, Ashot Margaryan, Helene Wilhelmson, Constanza de la Fuente Castro, Morten Allentoft, Peter de Barros Damgaard, Anders Johannes Hansen, Sofie Holtsmark Nielsen, Lisa Mariann Strand, Jan Bill, Alexandra Buzhilova, Tamara Pushkina, Ceri Falys, Valeri Khartanovich, Vyacheslav Moiseyev, Marie Louise Schjellerup Jørkov, Palle Østergaard Sørensen, Yvonne Magnusson, Ingrid Gustin, Hannes Schroeder, Gerd Sutter, Geoffrey L. Smith, Christian Drosten, Ron A. M. Fouchier, Derek J. Smith, Eske Willerslev, Terry C. Jones, Martin Sikora.

Correspondence to: Eske Willerslev (ewillerslev@snm.ku.dk), Terry C. Jones (tcj25@cam.ac.uk), Martin Sikora (martin.sikora@snm.ku.dk).

This PDF file includes:

Supplementary Text

Datasets and accession numbers of reference sequences

Ancient variola virus sequences from this study
Consensus sequences for the four higher-coverage ancient sequences aVARV-VK281, aVARV-VK382, aVARV-VK388, and aVARV-VK470 are available at the European Nucleotide Archive (https://www.ebi.ac.uk/ena) under project number PRJEB38129.

Dataset 1 (Human tree)
The four ancient sequences listed above, plus the following 51 VARV sequences:
L22579
DQ437580
DQ437581
DQ437582
DQ437583
DQ437584
DQ437585
DQ437586
DQ437587
DQ437588
DQ437589
DQ437590
DQ437591
DQ437592
DQ441416
DQ441417
DQ441418
DQ441419
DQ441420
DQ441421
DQ441423
DQ441424
DQ441425
DQ441426
DQ441427
DQ441428
DQ441429
DQ441430
DQ441431
DQ441432
DQ441433
DQ441434
DQ441435
DQ441436

DQ441437
DQ441438
DQ441439
DQ441440
DQ441441
DQ441442
DQ441443
DQ441444
DQ441445
DQ441446
DQ441447
DQ441448
Y16780
NC_001611
V563: LT706528
V1588: LT706529
VARV-VD21: BK010317

Dataset 2 (Full ML tree)
Dataset 1 plus the following 33 sequences (abbreviations: TATV: Taterapox virus, CMLV: Camelpox virus, ECTV: Ectromelia virus, CPXV-Gri: Cowpox virus Gri, CPXV-Ger: Cowpox virus Germany, CPXV-BR: Cowpox virus Brighton Red, VACV: Vaccinia virus, VACV-Ank: Vaccinia virus Ankara, VACV-Cop: Vaccinia virus Copenhagen, MPXV-Zai: Monkeypox virus Zaire, HPXV: Horsepox virus):

TATV: DQ437594 (NC_008291)
CMLV M-96: AF438165 (NC_003391)
CMLV-CMS: AY009089
ECTV: NC_004105
ECTV: KY554976
ECTV: JQ410350
ECTV: KJ563295
CPXV: KC813508
CPXV-Gri: X94355
VACV-Ank: AM501482
CPXV: HQ420893
VACV: AY313848
VACV: KF179385
CPXV: DQ437593
VACV: KJ125439
CPXV: KC813493
VACV: JX489138
CPXV: HQ420895
CPXV: KC813509
CPXV: HQ420897
CPXV: KC813492

CPXV: HQ420894
CPXV: KC813511
CPXV: NC_003663
CPXV: HQ420899
CPXV: HQ420896
MPXV-ZAI: NC_003310
CMLV: KP768318
HPXV: DQ792504
MPXV: DQ011156
Akhmeta: MH607143
VACV: NC_006998
VACV-COP: M35027

Dataset 3 (References for analysis of gene-inactivating mutations)
17 sequences, as described in Hendrickson et al., 2010 (6):
CMLV: AF438165 (NC_003391)
CPXV-BR: AF482758 (NC_003663)
CPXV-Ger: DQ437593
CPXV-Gri: X94355
ECTV: AF012825 (NC_004105)
MPXV-WR: AY603973
MPXV-ZAI: AF380138 (NC_003310)
TATV: DQ437594 (NC_008291)
HSPV: DQ792504
RPXV: AY484669
VACV-MVA: U94848
VACV-Cop: M35027
VACV-WR: AY243312 (NC_006998)
VARV-BRZ: DQ441419
VARV-KUW: DQ441433
VARV-SLN: DQ441437
VARV-SAF: DQ441436

Site descriptions

All sample materials described below are located at and curated by the Lundbeck Foundation GeoGenetics Center, Globe Institute, University of Copenhagen under the supervision of Dr. Lasse Vinner.

VK388: Lødingen, Nordland, Nordland 253 (Lisa Mariann Strand, Jan Bill)

The skeleton identified as Schreiner number 253 was contextually located in Lødingen, a municipality in Nordland county. The burial, covered by a small mound, yielded an ornamented tool of reindeer antler as the only datable grave goods. The ornaments, typical Samic in character, and the grave form dates the burial to the late 6th to 10th century. Two sigma 14C dating, uncorrected for possible marine reservoir effects, gives an age range of 603-653 CE. The skeleton is incomplete, some facial and post-cranial bones are unaccounted for, such as wrist and ankle bones. This could be explained both by the excavation method, and by different management strategies since 1889, when the skeleton was found. Measuring stature was not feasible for several reasons such as non-closing epiphyses and post-mortem erosion. The age at death was between 12–17 years, leaning towards the later spectrum in this ratio, due to the evaluation of epiphyseal closure on the long bones, spine, pelvis and the general tooth wear (*82*).

The skeleton exhibited signs of healed antemortem trauma on several rib bones, and one possible at one calf bone. The uncertainty of the latter is based upon the grade of post-mortem erosion and bone fragmentation. The skeleton also shows signs of active perimortem infections and inflammations (Table S9, Fig. S21), and Cribra orbitalia, which indicates anemia, and thus vitamin deficiency (*83*, *84*). Isotopic analysis shows that the individual was not born in the area and that he experienced a marked dietary shift during his life (*85*).

VK255 and VK470: Gnezdovo, Russia (Tamara Pushkina)

The Gnezdovo complex of archeological monuments (main period of its existence covers the early 10th to the early 11th centuries) is situated about 13 km westward of present-day Smolensk. The complex occupies a vast territory on both banks of the Dnieper river and consists of two settlements and seven groups of burial mounds that surround the settlements (*86*). The complex includes 13 coin and coin-and-items hoards of the 10th century. The hoards were found mainly in the territory of the Central settlement, the principal and the most explored settlement (*87*).

At the present time, the Gnezdovo settlement that belongs to the early 10th to 11th century is designated as the only early town centre in the Smolensk part of the Dnieper basin and the biggest monuments of the Old Russian state emergence in the territory of eastern Europe. Many characteristics obtained in the course of archeological study link Gnezdovo with town centres that existed in northern Europe in the late first-early second millennium CE, such as Birka, Hedeby, etc.

The monument's scope and its role in many respects were determined by Gnezdovo's geographical position at the junction of eastern Europe's river systems that connect the Baltic with the Black Sea and the Caspian Sea.

The choice of the place for settlement made by Slavs and Scandinavians who arrived there practically simultaneously was determined by the landscape situation and favorable climatic conditions at the end of the first millennium CE: at that relatively dry-weather period the Dnieper flood plain was not drowned during high waters and was suitable for reclamation. Initially, the settlement territory was inconsiderable and amounted to no more than 1–2 hectares but as early as the first stage of the settlements' existence, the Central settlement, i.e., the fortified part of the settlement was formed. A small number of the burials (about 1% of the collection) belong to the initial stage of the settlements' existence. This number of burials completely agrees with the monument's initially modest scope.

The heyday of Gnezdovo was in the second half of the 10th century: the settlement area increased and attained its maximum (about 30 hectares) due to expansion of built up area within the floodplain limits and development of considerable new area along the right bank of the Dnieper above the floodplain and in inward lands.

The majority of the ~1000 burial mounds explored in all five groups of mounds were built during the middle and the second half of the 10th century. Rare burial mounds with cremations (such mounds were built near the eastern, north-western, and western boundaries of the settlement) are related to the earliest period of the settlements' existence. Some mounds of the Forest mound group and partially the Central mound group belong to such burials. The well-known mound L-13 is one of these mounds. The L-13 mound gave the earliest find in Gnezdovo, a Byzantine amphora with the Slavic inscription-graffito. In the middle and the second half of the 10th century the other mound groups (the Dnieper, the Ol'shanskaya, and the Right-Bank Ol'shanskaya groups) were formed. These late mounds are situated along the Dnieper bank downstream. At the same time the Forest and the Central mound groups continued to grow.

Single and paired burials performed in accordance with cremation rites under low mounds and so-called big mounds with rather complicated cremation rites as well as single inhumations in graves and chambers under mounds belong to the late period of the settlement existence. Male burials in chambers accompanied with riding horses belong to the same period. By the end of the late period, the total number of burial mounds attained 4500–5000. Cremation is the predominant burial rite in Gnezdovo mounds (cremations comprise 52% of burials), while inhumations compose about one third (31%) of burials in the collection of the explored burials. The remaining 17% are so-called empty mounds, where remnants of burials have not been detected.

As judged by materials obtained in the mound necropolis, the population of Gnezdovo attained approximately 800–1100 persons and this population distinguished itself by a stable demographic structure and obvious social differentiation.

Exploration of the settlement territory has demonstrated its variegated structure and availability of different zones: residential and manufacturing zones as well as related areas, with Gnezdovo functioning as a river port. Traces of manufacturing activities represented by findings of tools, rejected products, intermediate products, and raw material stocks have been noted on all explored plots, from the earliest period of the settlement existence and onwards.

At the heyday of the settlement, no less than five manufacturing zones functioned in Gnezdovo. Findings related to processing of iron, nonferrous, and precious metals are concentrated in these zones. Workshops produced adornments for Scandinavian and Slavic women, bridle and belt plates made in the Volga Bulgarian tradition, and items specific for the local culture of long mounds. The topography of Oriental coin silver findings and Scandinavian antiquities witness their concentration in Gnezdovo and its immediate vicinity.

About 450 Oriental, Byzantine, and west European coins (the majority of these coins are fragmented) have been found in burials and the occupation layer. Dates of coin minting vary from the 6th to the mid-11th centuries. Thirteen Gnezdovo hoardings contain over 1400 Oriental silver coins (intact and fragmented, with punches and riveted eyelets), two Byzantine *nomismae*, weighing scales, and plummets as well as various adornments. Adornments include 190 gold and silver items. There are items of Slavic, Scandinavian, and Oriental origin among these adornments.

If judged by concentration of Oriental silver, Gnezdovo is comparable with north European monuments such as Birka and Old Russian monuments such as Kiev, Rurikovo gorodishche (i.e., a place of abandoned old fortified settlement), and Novgorod (the earliest layers). So-called Byzantine imported goods (glazed pottery, precious textiles etc.) comprise a considerable part of findings. Abundance of Scandinavian items (adornments, items of everyday use, amulets, etc.) is one of the most expressive features of the Gnezdovo material culture. These findings are matched with data of burial rite analysis of which allows the conclusion that at least 25% of burials are Scandinavian.

Data from sporo-pollen, carpological, and osteological analysis bear witness that cultivated plants and pastoral farming played an important role in the settlement economy and to some extent provided for vital requirements of the settlement inhabitants. The inconsiderable areas of arable land discovered in the western part of the Gnezdovo complex beneath mounds of the Dnieper group and the character of the landscape at that time speak of the inability of the early town population to provide themselves with agricultural produce and the need to obtain a considerable fraction such products from elsewhere.

The youngest burials of the Gnezdovo necropolis are few and can be dated to the turn of the 10th and 11th centuries. Archeological materials indicate that a rather active life continued in Gnezdovo in the early 11th century, but that Gnezdovo gradually lost its unique character of an urban centre and became a feudal estate. The final stage connected with the town extinction and transfer of its functions to Smolensk, as described in chronicles. The fate of Gnezdovo is similar to that of certain other early town centres of northern Europe. It is believed that these centres ceased to exist for a variety of reasons. These include the piecemeal termination of Islamic coin importation during the decade of 970 CE, change of trade routes, and a conflict between the local elite and the central power of the emerging state (*87*).

Burials (mound 212 and mound 292). Both burials were made under mounds in the same burial site; the distance between the mounds is ~150 m.

Mound 212 (VK470): By anthropological data, these are the remains of a female, about 25–35 years of age (aging and sexing was done by D. Pezhemsky). The inhumation was in a cell, the position of the body, possibly, was sitting; the head was to the west. A set of costume jewelry, typical of a female Scandinavian, was found in the burial. Thus, the burial could be Scandinavian. The date of the burial is the second half of the 10th century.

Mound 292 (VK255): By anthropological data, these are the remains of a female, about 20–25 years of age (aging and sexing was done by D. Pezhemsky). The inhumation was in a pit, the body was on its back with the head was to the west. There were few jewels, one being a ring, which was a decoration of the temporal area of the head (typical for Slavic). Thus, the burial, possibly, was Slavic. The date of the burial is the second half of the 10th century.

VK382 (sample 1132), SHM 22231, Nabberör, Böda parish, Öland, Sweden (Helene Wilhelmson)

The site: The boat burial 'Nabberör' is a solitary burial site located close to the beach of the protected northern cove, and natural harbour (Grankullaviken) in the large Baltic island of Öland. It is dated by artefact typology to the Late Vendel or Early Viking Age (700–800 CE) and was an unburned boat covered by a stone cairn. Two sigma 14C dating, uncorrected for possible marine reservoir effects, gives an age range of 640–770 CE. There are four cremation boat burials excavated on the island. They are far from Nabberör in the central part of the island and in the context of large burial grounds (*88–92*). Nabberör is the only boat burial with uncremated remains and contains a large number of human remains of varied ages (minimum 8–11 individuals, see details below) as well as many animal remains and various artefacts. The burial was large – the cairn measured at least 15 m in diameter and 2 m high. The boat was approximately 10 m long and 3 m wide (*93*).

Nabberör is unusual in the regional context of Vendel and Viking age boat burials in the Baltic area (Sweden, Finland, Russia) and North Atlantic (Denmark, Norway, England, Scotland). The 'standard' boat burial from this period usually includes one or two people in a wooden boat with artefacts as well as animals in and around the boat (*88*). The number of domestic animal companions (dogs, birds, sheep/goat, pig and horse) are usually one per species, although there are variations. Nabberör is truly very different in these respects as it contains many animal individuals of all domestications as well as many humans. The number of human remains in one boat is rivalled only by the Salme (Saarema) site in Estonia where more than 30 adult individuals, the dead from a single battle event, were found in two boats (*94*).

The investigation 1938-39: The central part of the Nabberör burial was looted in an event preceding any documentation and records of the stone cairn. Here we argue that the looting took place in the early 1800s due to a coin dated to 1805 found in the excavations (*93*). In a photo taken of the cairn in 1928, the looting pit is clearly visible (*92*). In the early 1930s further destruction almost removed the cairn entirely as the stones were sourced for building materials. There are no records of any finds made at that point, however, the locals did bring the cairn to antiquarian attention in an inventory of ancient monuments some years later, in 1938.

In the rescue excavation in September 1938, archaeologists documented the extent to which the site was manipulated by the looting. The looting was restricted to the centre of the grave and to their surprise the material from the looting pit yielded several objects of great value, for example a decorated sword and shield with silver inlay on gold, as well as commingled human remains. They also fully excavated the undisturbed part of the burial, finding over 200 iron rivets from the hull, the remains of more humans and animals, and additional artefacts (*93*, *95*). After the excavation the area was evened out and no traces of the cairn or burial remain today. The excavation is well documented for its time, with plans and photos as well as detailed observations. The finds were collected, including bones, and are curated in the national museum.

The burial was processed by Anderbjörk and presented in both an archive report (*93*) and a popular article (*95*). Both texts present the highly furnished boat burial with a few individuals and animal companions. Anderbjörk (*95*) was hesitant to compare the burial with similar contemporary finds (for example Vendel and Birka, boat and chamber graves) and concluded that it was unusual and odd in its characteristics. However, there was no interest in further study of any aspect of the Nabberör burial until 2011.

Investigations from 2011 and to the present: Some of the human and dog remains from Nabberör were investigated in 2011 as part of a thesis (*96*) and underwent stable isotope analysis. Since 2016, the entire find is being analysed – all skeletal remains (animal and human), artefacts as well as the documentation of the 1938 excavation – by a team from Lund university (Helene Wilhelmson, Kristina Jennbert, Elisabeth Iregren, Stella Macheridis, and Fredrik Ekengren). Some of the results of this project are presented below for the first time, the work of Wilhelmson for the human remains.

Stable isotopes: Two humans and one dog have been analysed for $^{87}Sr/^{86}Sr$, $\delta^{18}O$, and $\delta^{13}C$ in enamel apatite (*97*). These results were interpreted as one of the individuals and the dog being clearly non-local, non-native, to Öland. The human and dog found together (in A4) did not originate from the same place as they had very different Sr values. They all had very similar isotope values, $\delta^{13}C$ and $\delta^{15}N$ in bone collagen, indicating a similar diet for the humans and the dog (*96*, *98*) (Table S10).

Demography: Among the dead in Nabberör is a child aged 6 years, and at least two more children under the age of 15 years (cf Table S11), as well as juveniles/young adults and adults. The bones were too poorly preserved, or lacking suitable elements to yield an osteological sex estimations except for one individual (1150, A10), a possible male. The analysis of the four teeth samples that yielded preserved aDNA (*99*) showed all these individuals to be male (Table S12).

The number of human remains in Nabberör is at least 8–11 persons (Table S11). The preservation was poor in the sense that primarily teeth, which are more resilient than bone in most environments, were available. In the undisturbed parts of the boat other skeletal elements were found making up one or more complete or fragmented individuals. Looting is no doubt responsible for the great loss of skeletal elements, as well as general preservation due to the later disturbance of the topmost cairn stones.

The total MNI (minimum number of individuals) represented in the skeletal remains needs to be considered carefully. A conservative approach is to consider the entire boat as one assemblage, not taking into account the observations of the spatial distribution noted by the excavators. However, for the most part, their assemblage definitions correspond to one more-or-less complete individual (with the exception of A7 which is the looting pit).

The conservative MNI: This MNI is entirely based on teeth. They were divided into individuals based on duplicates of a single tooth number (for example FDI 46) in addition to:

- symmetry in morphology and size (for example FD 46 and 36)
- development (maturity of the tooth; i.e., separating children (with not fully developed permanent teeth or with deciduous teeth) from adults with completely formed teeth
- attrition
- preservation (colour etc.)

This gives an MNI of eight for the boat as one unit (A4 (1 individual) A8 (1 individual) and A7 (6 individuals)) and only includes the human remains found with teeth.

The assemblages included MNI: If we consider each assemblage (‘anläggning’) to be separate units – as they indeed were perceived to be while excavating – then there are five individuals in the peripheral areas of the boat (A4, A8, A10, A11) and an additional six in the centre looting pit (A7). This MNI is 11 people.

The true number of interred individuals could of course be higher due to the obvious considerable taphonomic loss of skeletal remains resulting from looting.

For the majority of the individuals there was no indication of a possible cause of death, as expected due to the high level of fragmentation of the remains. One individual (an adult, 1150) was however most likely murdered, showing a perimortem sharp-force trauma to the skull.

VK281: Bårse 'Tyvestensagern', Museum Sydøstdanmark SMV 28/86, AS 1/1987, Grave A (Marie Louise Schjellerup Jørkov and Palle Østergaard Sørensen)

In 1986, prior to the construction of a bridge and a new section of the motorway A20 between Rønnede and Udby on Zealand (Denmark), several large archaeological investigations were completed at a small village named Bårse (See Fig. 1 in Sørensen 1995 (*100*)). The excavation sites lay along a sandy ridge on the northern side of a small stream named 'Risby Å'. A so-called 'causewayed enclosure' from the early Neolithic period was found at 'Markildegård' on the northeastern slopes of the sandy ridge. Above the Neolithic site, on the top of the ridge, a cemetery from the Viking age was excavated (See Fig. 3 in Sørensen 1995 (*100*)). This cemetery, called 'Tyvestensageren' (Bårse Sogn, Bårse Herred, Præstø Amt, Sb67), consisted of eleven graves and five features which did not contain human remains, but that were also considered to be graves due to the shape and orientation of the features.

Grave A was 2.80 m long and 2 m wide, and was by far the largest grave investigated at the site. Three large stones were placed on the top of the grave. The grave was at first investigated to a depth of 80 cm without any finds, and was at that time considered to be empty. Sometime later, construction workers took a sample of the gravel from the bottom of the grave, and hit a layer of charcoal. The grave was then investigated to a depth of 1.15 m, and a human cranium was found. The rest of the skeleton was found a little deeper, which suggests that the head of the deceased had been placed on a podium or a thick pillow. The body was on its back with the right arm across the stomach. The bottom of the grave was reached at a depth of 1.60 m. The deceased was placed with the head to the south-west, in contrast to all the other graves at the site, where the body was orientated with the head to the north. The very large grave and the unusual orientation of the body seems to indicate that the person in grave A was more important than a normal person.

Besides a few iron nails and a small whetstone, which were found in the upper fill, the only grave goods from grave A were a small belt buckle and a belt bracket, both made of bronze. These items were found in the pelvis region of the body. The finds from grave A, and a fibula found in another grave, are the only dateable items from the cemetery, and they date the graves to the 10th century.

Only few pathological observations could be made on the skeleton from grave A. They included porosity in the palate and diffused areas of pitting on the ectocranial surface of the sphenoid bone. An upper thoracic vertebrae showed a Schmorl's node lesion (due to necrosis of the vertebral disc). The dentition showed slight calculus deposits. There was agenesis of the lower left second premolar and the second left deciduous molar retained in its position. The upper central incisors showed horizontal grooves on the labial side of the crown due to abrasion (Fig. S22).

Around 400 m northwest of the cemetery at 'Tyvestensageren', an excavation in 1973–76 revealed a 70 m long stretch of road, which crosses the Risby Å at this point. The road was partly built of stone and wood. Under the road, parts of a sledge and a wagon wheel, both made of wood were found (See Fig. 4 in Sørensen 1994 (*101*)). This road is also dated to the 10th century.

Tyvestensageren: 686955.59, 6113106.00 UTM euref89
The Risby road: 686699.85, 6113507.07 UTM euref89

VK168: Sample SK1899 (Ceri Falys)

The body of SK1899, found in the middle of a mass grave at St. John's College, Oxford, was heavily truncated by a pipe-trench. Only portions of his upper body were present for osteological analysis. The individual was male, and was estimated to be aged between 16 and 25 years at the time of his death. The mass grave contained the remains of at least 35 male individuals. As with the majority of men in the mass grave, SK1899 showed several perimortem blade wounds to the back of his upper body (one to the back of his head, and multiple to his ribs and left shoulder blade). Many of the injuries were inflicted from behind, and no defensive wounds were observed on the skeletons. Several skeletal elements also displayed charring. The only other pathological alteration to his remains was active new bone formation on his left zygomatic bone (left cheek), although a differential diagnosis could not be suggested as the majority of his skeleton was not present for analysis. Carbon dating of our mass grave was problematic, due to the marine reservoir effect. The best estimate for the formation of the mass grave is 880–1000 CE. The osteological findings and historical documentation indicate that these men were the Danish victims of King Aethelred's decree ordering their extermination in 1002 CE, during the event known as the St Brice's Day Massacre (*102*).

VK443: Sample 1101, SMH 21367, Grave 24 (Helene Wilhelmson)

The skeleton sampled as VK443 was buried in northern Öland. It is burial number 24 in the grave field known as 'Forgalla skepp' (Forgalla ship) id 21367 (SHM); Böda parish; RAÄ5:1.

The burial is a combination, or possibly a creolization, of burial rites and similar burials are found in Viking age Öland (cf discussion in Wilhelmson 2017 (*96*)). The body was oriented in a Christian burial tradition (West-East) contrary to the pre-Christian tradition (going back at least to 500 BC) with a North-South orientation. Supine (on the back) burials of both orientations are found throughout the Viking Age on Öland and were probably practiced in parallel. The non-Christian elements of this burial are significant. Firstly, the pit with the body was covered by a stone paving (also oriented West-East) and a large circular cairn of stones (7 m diameter and 0.75 m in height). Secondly, several personal artefacts were found with the skeleton. A comb was found by the left arm/chest, a knife on the sternum, and a strike-a-light on the neck. The creolization of burial rites is apparent as there are some highly local elements for Öland (the use of local limestone slabs for a stone paving as well as a cairn) and some possibly specifically Slavic and Christian features that are discussed below.

The Christian orientation (West-East) and pre-Christian rite of placing everyday objects in the grave are defining characteristics of the 11th century (possibly contemporary to VK443) burial ground Ndr. Grødbygård in Bornholm (Naum 2008; Price et al. 2012:95f (*103*, *104*)). Bornholm is another large island in the Baltic south of Öland. The Ndr. Grødbygård burial ground is interpreted as an early Christian Slavic cemetery (Naum 2008 and discussion in Price et. al 2012 (*103*, *104*)). Another interesting similarity is the positioning of the hands over the pelvis (Fig. S23) in VK443, an unusual position in Viking age/Early medieval Scandinavia. This

has been found in at least one burial (Grave 210; Price et al. 2012: 96 (*103*)) in Ndr. Grødbygård. The skeletal remains in most burials were too poorly preserved to indicate anything about the hand positioning and how frequent it might be. In conclusion, despite that none of the artefacts with VK443 are of a typical Slavic type the similarities in burial practice with the Slavic burial ground in Ndr. Grødbygård are significant and suggest a Slavic element in the burial ritual for VK443.

The artefacts date widely to the Viking Age (800-1050 AD). Isotopic analysis of this individual shows he was a non-local to Öland (*96*). The Sr value (0.7319) is similar for example to middle Sweden and the Uppsala region (*105*) or Birka (*106*). The value is also similar to values found in individuals buried in the Salme Boat burials (*94*), suggested to have originated in the Uppsala/Mälaren region from both Sr values and associated artefacts. Similar Sr values have also been reported as local bioavailable baselines in Finland (review in (*106*):31). According to a recent aDNA analysis of this individual his ancestry was estimated as primarily Finland (43% contribution), Poland (30% contribution), and Sweden (24% contribution) (*99*). The Sr value for VK443 would fit well with origins in Finland or Middle Sweden (Mälaren/Uppsala region) but not with Poland (compared to baseline in (*103*):103). Genetic heritage and place of residence in childhood (indicated by Sr) are two different things and cultural affiliation (as a Slav in this case) can be yet another. It is interesting nonetheless that all available sources (burial, DNA, Sr) all seem to indicate a mix of both geographical and cultural heritage.

The skeleton was well preserved, and most bones were present except for some of the smallest. Bone fragments from at least one additional individual were also found. The sex was estimated to be male, based on skull and pelvis. Living stature was estimated as 174.9 ±4.5 cm (*96*), an average height for a man in Viking age Scandinavia. The age of death is estimated to 20-23 years old taking both pubic symphis (Suchey Brooks phase 2; 23,4±3,6 years; 19-34 years 95% range; (*107*)) and epiphyseal fusion ((*108*); Table S13) into account. However, if the pubic symphysis is disregarded, the individual could be slightly younger, 17–23 years old. The bones displayed extensive pathological bone remodelling, resulting from persisting disease. Overall, the changes can be interpreted as a case of pulmonary osteoarthropathy i.e., lung cancer (*96*):452. There was a small depression in the frontal bone on the left side, likely a healed trauma. A rib, also on the left side, had a healed partial fracture.

VK533: Sample 1076, SHM 28364, Grave A136 (Helene Wilhelmson, Ingrid Gustin)

The site: The large grave field Sörby-Störlinge was excavated in 1964–66 as part of Ulf Erik Hagbergs research of Iron Age burials in Öland. The grave field was intensively used during the Iron Age and in one instance a single stone cist was reused adding more proper burials on at least three occasions (*109*, *110*). Today it has been established with radiocarbon dating that these three individuals, found in the same cist, span almost 1000 years from the Pre Roman Iron Age to the Viking Age (conventional dates 2065±45 and 1175±50) (*96*). Both inhumations in various forms as well as cremations, all likely Iron Age, have been found in this grave field.

Burial A136: The inhumation burial A136 is described as covered by a stone pavement relating to a later grave (A128) also including a cremation pit. It was a pit burial with the body placed in a supine position. The orientation of the burial is unusual for Early Iron Age Öland, it is West-East instead of North-South.The burial is described in Hagberg (1966) (*109*), Beskow-Sjöberg & Arnell (1987) (*89*) and in Wilhelmson (2017a) (*96*). Before going into detail the date

of the burial needs to be addressed as this is today considered to be erroneous in the earlier publications.

Date of the burial, local (grave field) and regional (island); discussion of the tradition of East-West oriented burials on Öland: The burial, id 1076, A136 in the Sörby-Störlinge grave field was dated to 150/160-260/270 CE, (period V:1, Late Roman Iron Age (*109*, *110*). This date should today, in my opinion, be revised to Viking Age, 800–1050 CE. The reasons for this revision are outlined in detail below.

A recent study surveyed all inhumation burials on Öland, with skeletal remains intact and available for analysis, using an interdisciplinary bioarchaeological approach (*96*). 109 human remains from the entire Iron Age were selected and studied with isotope analysis (including 14C). The results from the 14C dating showed that many assumptions regarding the dating by grave form (shape, structural elements etc.), and/or adjacent graves, when artefacts were lacking, to be erroneous (*96*, *98*). They also confirmed dates indicated by artefact typology (details in Wilhelmson 2017b (*98*)). Further, this study also proved that the use of most single burial grounds to span almost the entire 1550 years of the Iron Age (500 BCE–1050 CE), and not primarily Roman Iron Age, as previously suggested by Hagberg in 1979 (*111*). The number of confirmed Viking age graves in Öland increased greatly as a result of this study (*96*, *98*).

There are in total 14 inhumation burials oriented East-West (included in (*96*), see details in Table S14) available from Öland. There are more East-West burials on Öland that are not reported here, as these have not been possible to study in detail to perform, for example 14C dating, due to lack of bones. As mentioned above, the 14C dataset presented in this study (*98*) showed much greater complexity in burial chronology, grave field continuity and relative dating in Öland than previously assumed.

I argue that all of the 14 cases of East-West burials (Table S14) should be dated to the Viking Age, with one possible exception. The exception is the burial of a partly burned articulated individual, dated to the Pre Roman period by 14C, id 1098. The pattern of burning suggests something very different from a proper cremation and the body was also articulated and, likely clothed. A fatality resulting from a burning house is one possible explanation. The radiocarbon analysis for this individual is however from the 1970s, and from a partially burned individual with potential collagen damage, so there is a possibility that the results would be different using today's methodology for radiocarbon dating. The date could nonetheless be correct and this anomaly in orientation may well be explained by the unusual circumstances of the death of this individual.

Today, 12 burials with East-West orientation found on Öland are dated by specific typology and/or 14C (Table S14) to the Viking Age. The East-West orientation of a grave is possibly an early Christian influence, thus not likely to appear long before 800 CE in Scandinavia. The other elements of the burial above ground (cairn, stone pavement etc.) as well as below (cist, wood coffin or a simple pit) vary within this group. This could be indicative of regional traditions (for example, a coffin if the deceased had a southern cultural affiliation and a 'traditional' Ölandic stone cist in the new East-West orientation if born on Öland), or chronology possibly, but still within the Christian burial practice if the East-West orientation was used (cf. discussion in Naum 2008 for the Baltic island of Bornholm). In general for Öland, today, with all the new 14C dates (Table S14) in mind, a burial with East-West orientation should be assumed to most likely be from the Viking Age.

The reason for the much earlier date of A136, id 1076, in the literature (*109*, *110*) is the fibula found in the upper grave fill. This was suspected to be a fragment belonging to the needle

and spiral found further down in the grave and on the bones of id 1076. But, it is also possible that the fibula is redeposited material (from an earlier destroyed grave) as it is not directly related to the skeleton. This is an intensely used burial ground for more than 1000 years. The knife found on the skeleton (Fig. S24) is of a general type and may be Late Iron Age as well as Early Iron Age. Therefore I argue that the date based on artefacts is not conclusive and only indicates Iron Age in general. Locally, on the grave field Sörby-Störlinge, there is a second East-West oriented burial (id 1058, A134) which was dated to the Iron Age/Early Iron Age from burial form. This burial is now dated by 14C to the Viking Age ((*98*), Table S14). Therefore, id 1076 (A136), being the only other East-West burial in this specific grave field should also be considered most likely Viking age. As for burials in North-South orientation in Sörby-Störlinge, some of these have also been concluded to be Viking age, based on 14C (Table S15). This demonstrates that the presence of Viking age burials in the grave field is much more prominent than previously assumed. The reasons for these assumptions can only be speculated. It may relate to a lack of 14C dates (not available at the time of excavation) and few clearly Viking age artefacts found in the graves. Possibly, it could also be a reflection of Hagberg's research portraying the Early Iron Age, in contrast to the Viking Age, as *the* golden era of Öland in his research including key sites in Öland from 1960 to the 1980s.

The knife: The only artefact that is directly associated with the skeleton (id 1076) in grave A136 is a knife resting on the individual's right clavicle (Fig. S24). The date of the knife, and the burial at large, is argued here to differ significantly from the previously published date of the burial. Therefore, the basis for a date of the knife needs to be elaborated in some detail. The date presented below is based on:

- a study of dimension and typology. This is based on the photo (Fig. S24), drawings, and details located in the archives, as well as via the curating museums' digital archive. The current location of the actual knife, as well as many other iron artefacts (if still in a museum collection) is unknown today. It is possible it has not survived in storage, being made of iron and collected over 40 years ago.
- a comparison to relevant similar knives found in burial contexts (also based on photos, drawings, etc., available in the literature).

Description

The knife was described as having a total length of 10.6 cm. The handle still had some fragments of wood shafting when first discovered. The blade was 1.06 cm wide with a triangular cross section. The back of the blade was 0.5 cm across. From the photo (Fig. S24) it appears the back and the cutting edge are straight. The handle appears to be about half the length of the entire knife, i.e., blade and handle are equally proportioned. The base of the handle is 1.1 cm and it tapers back to the end. The handle had a rectangular cross section.

There was another knife of similar size (11.5 x 1.1 cm) but different shape in grave A164 in the same grave field as A136. Grave A164 was oriented North-South and a limestone cist and the 14C (of the bones) gave a date in the Viking Age (Table S15). A164 was only dated Iron Age in general before the 14C dating. This is not conclusive evidence of the date of the knife in A136 but it does lend support to a date in the Viking Age.

Typology and parallels: There is a lack of a comprehensive and detailed study on Iron Age knife typology in Sweden. The use of a typology seriation in dating an individual knife is therefore problematic on many levels. Regional variation as well as chronological variation are poorly understood today. Knives are also in a sense a highly functional item and may as such be variable within the same time period. If the intended use of the knife was primarily a single

specific purpose there are likely many variations due to customization. That said, there is one case study that is of relevance to the knife in this specific burial.

Helgö and Birka are located in Mälardalen, central Sweden, and knives have been found in the excavations of both graves and settlements there. These sites are of relevance also from a regional perspective as the individual in the grave (id 1076) might well have originated from the Mälardalen region (c.f. discussion on Sr below).

The knives found in Helgö and Birka (in Mälardalen, central Sweden) respectively have been compared (*112*). It is therefore possible to see how an individual knife holds up to some general observations made for the knives from Helgö – a site representing a time span from 200-800 CE – and the ones from the trading place Birka, representing the time period 750–970 CE. It is still very problematic to give a specific date for a single knife. However, we can determine what is a more (and a less) likely date. Arrhenius found that 70% of the knives in Helgö were shorter than 7.0 cm, while in Birka this was true only for 5% (in the graves) and 10% (in the settlement) (*112*). Another distinguishing feature was that the knives from Helgö primarily had concave edges, a quality usually associated with heavy use and repeated re-sharpening. They also more commonly had a curved ridge than the knives in Birka. The most frequent knife shape in Birka, found in 50% of knives, was both a straight ridge and a straight edge (*112*).

In the light of the Birka and Helgö material, the knife from A136 is more likely Viking age (in concurrence with the common types found in Birka) than from previous periods of the Iron Age, including Roman Iron Age (0–400 CE). The straight ridge and edge are features that show that a dating to the Viking Age is possible, but it is at the same time not conclusive enough to determine the knife to be an exclusively Viking age type. The burial needs to be dated by other means as well, such as grave form typology (which above was concluded as most likely indicating Viking Age), and a radiocarbon analysis could be the only conclusive solution. It is notable that there is no reason to date this burial to the Roman Iron Age today and at least two aspects (the orientation and the knife) support a date in the Viking Age.

The buried individual, id 1076, in grave A136: This burial was included in a recent study of Iron Age burials in Öland using current bioarchaeological methodology, the results of which are presented below for the first time. The isotopic results were published before (*96–98*). A second individual was found amongst the skeletal remains, likely a remnant of an earlier burial as this was an intensely used burial ground during the Iron Age, as mentioned above. The second individual was fragmented, only two cervical vertebrae, including the atlas (first cervical vertebrae), from a much larger individual. From the archive documentation of the burial (*109*) it is clear that the left hand was positioned on the right hip and that the skeleton appears well articulated – as expected from a pit burial. Despite the find of a broken fibula in the fill and partly on the body the skeleton appears undisturbed anatomically, so if the grave was reopened this happened when the body was still fresh. The osteological sex of the skeleton is female, aged 50–60 years, but the genetic sex is male. Her stature is calculated to 160.8±4.49 cm. She has a healed fracture on her right third metacarpal, palmar side of diaphysis. This resulted in a shortening of the bone by 0.2–0.4 mm. Further, she suffered from bilateral rotator cuff syndrome (diagnostic definition in Waldron 2008 (*113*)), i.e., impairment in both right and left shoulders.

The $^{87}Sr/^{86}Sr$ value, obtained from premolar enamel, is 0.733748 (*97*). She is therefore clearly non-local, non-native to Öland and could be from, for example, northern Sweden or the Uppsala region (cf Bäckström & Price 2016 (*105*)) or another area further away.

Bone collagen samples resulted in values of -19.9 for $\delta^{13}C$ and 14.2 for $\delta^{15}N$. This reflects the adult diet in the decades before she died. These values are corresponding to contemporary

Late Iron Age individuals in Öland, only slightly higher in $\delta^{15}N$ (mean values for the period are $\delta^{13}C$ -19.4±0.5 and $\delta^{15}N$ 12.7±1.0). The diet isotopes likely reflect the local pastoralist subsistence in Öland. During this time it seems they were intensely utilizing meat and secondary products (milk), mostly from sheep, thus achieving the high $\delta^{15}N$ -values. As far as the diet isotope results are concerned, she appears to have lived on Öland as an adult (*98*).

FIN1: Sample CGG21429, 21429E, 6755-76 (Vyacheslav Moiseyev, Valeri Khartanovich)

The Varzino village was located on the right bank of the Varzin Bay (coordinates: 68.37, 38.38), not far from the coast of the Barents Sea.

Until the 1930s, Varzino village was used as a summer standing point by the Semiostrovsk group of the Kola Saami. After the transition of Saami people to a sedentary lifestyle, it became a year-round settlement. Because of the very isolated location of the settlement, difficulties with food supply, medical, and social care in the late 1950s, all inhabitants of Varzino were relocated to other settlements of the Kola Peninsula, mostly Lovozero (*114*).

The Varzino cemetery is located about 3 km from the village and is separated from it by a brook. Separation of cemeteries from the living space by water barriers is one of the characteristic features of the Saami culture. The cemetery was investigated in 1977 by an expedition of the Museum of Anthropology and Ethnography, Academy of Science, USSR under the guidance of I. I. Gokhman and V. I. Khartanovich. All excavated burials were dated to the end of the 18th and the end of the 19th centuries. Morphological study of the skull collection from Varzino revealed a high level of similarity between the Varzino population and other Saami groups of the Kola Peninsula (*115*).

The FIN1 sample is morphologically female, and 40–45 years old.

VK515, Nordland 4512 (Lisa Mariann Strand)

The skeleton identified as Schreiner number 4512, VK515, was contextually located in Bodø, a municipality in Nordland county, Norway. It comes from one of two burials, excavated 20 m apart, 70 m from the beach. Both burials provided sparse grave goods that dates them to 950-1025 CE. The difference between the contextual dating via grave goods and the 14C dating (664-768 CE) given in Table 1 is likely due to uncorrected marine reservoir effect in the latter. As VK515 is not a higher-coverage sample, it is not included in the regression analysis shown in Figure S14, so the uncertain dating has no effect there or on any other analysis herein. Isotopic analyses have revealed that VK515 grew up in another region than the one where he was buried, and that he possibly experienced a shift towards a more marine diet in the last part of his life (*85*). Mitochondrial DNA was identified as belonging to haplogroup H* (*116*).

The skeleton is incomplete and is heavily affected by post-mortem taphonomy changes such as dry breaks. The calvaria is joined together with steel thread. Measuring stature was not feasible for several reasons, such as post-mortem erosion and missing bone elements. The age at death was 18-22 years, based on the high degree of visibility of the iliac crest epiphyseal scar, degree of cranial suture closure, and dental wear (*82*).

One ulna deriving from the skeleton exhibits peri-mortem signs of being exposed to violent conflict, the same bone displays rust marks (Table S16, Fig. S25, panel A). Even though some peri-mortem activities have contributed to the condition of the bone, the trauma is still

identifiable, and caused the individual's death. The visible rust deposits could indicate the intentional placement of archaeological objects during the burial.

The skeleton exhibits signs of active peri-mortem inflammation before death on several long bones and carpals, among others on the tibia and fibula on both sides (Fig. S25, panel C). Schmorl's node (Fig. S25, panel B) was also identified in the lower region of the vertebral column, and there may be a connection to muscular stress markers identified in long bones from the appendicular skeleton and extensive labour (*117*). Other identifiable health signs are enamel defects observed on the teeth that were present. Enamel hypoplasia and an excessive amount of caries, which are present on the teeth, are two related conditions (*118*). All teeth connected to the skeleton displayed caries.

KHA1 K793 (17793E): Obdorsk' Khanty (Khalas-Pugor), Russia (Alexandra Buzhilova)

The site: The Khalas-Pugor site was investigated by T. D. Yanovich in 1909 (*119*). The site is a burial complex of a group of Khanty, who lived near the Obdorsk town (today Salekhard − administrative centre of Yamalo-Nenets Autonomous Okrug, Russia, straddling the Arctic circle). The Khanty are Taiga hunters and fishermen. They settled on riverbanks and used artificial fences (abatis), crossbows, and pits for hunting. The Khanty lived in dugouts and semi-dugouts of various designs, and later on in log huts.

Khalas-Pugor is characterized by uniformity of burial rites. Of the 154 burials, only two individual burials were placed in pits, the rest were above-ground. The burials present complex construction. Inside the wooden box-holmer was placed a half-boat, where the body lay, with feet to the bow of the boat. Sometimes the coffin was covered with another half-boat or planks of narts, birch barks from chums. Almost all the burials had a Southern orientation (*119*).

The funeral artefacts were located on three 'levels'. On the 'outer part' the artefacts were located near the burial in a leaning position or they were located on top of a box-holmer. These are skis, oars, bows, troughs, etc. The second 'level' was the space between the walls of the box-holmer and the half-boat on the coffin lid, where there are both large and small artefacts (such as idols). In the inner 'level', there were collections of 10 to 50 or more items close to the body. Almost all individuals had dishes, knives, axes, sometimes hunting weapons, metal products, and most frequently metal plates. In addition, miniature anthropomorphic images made of wood dressed in 'clothes' were found in the burials (these are idols, temporary containers of the deceased's soul).

Archaeological investigation of the artefacts, especially metal plates, allows the proposal that funeral complex was mostly used in the 19th century CE.

Burial 66: The remains of the child were located in a wooden box whose length did not exceed 100 cm. The width in the head area was 31 cm, and in the leg area about 29 cm. The head of the child lay to the South, the body was in an extended position, the arms were laid along the body, and hands were located under the pelvis. Both outside and inside the box were different burial artefacts. Near the box lay a snuff box and a spoon. In the box, to the left of the child's head, were beads and beads also lay on the chest. In addition, a spoon-cook lay near the right hand, a food-trough and 2 rings were located on the stomach, and a tablespoon was lying closer to the right leg (*119*).

In the anthropological collection (#6926) there were preserved only parts of the skull and mandible with few teeth. According to teeth development, the biological age of the child is about 3–4 years (*82*). Slight multifocal porous hypertrophic bone formation are recorded on the grate

wings of the sphenoid, and partly on temporal bones, on upper border of orbits, on maxilla near alveolar areas of frontal teeth and hard palate (Fig. S26). All the alterations are associated with infantile scurvy (Möller-Barlow disease) (*120*).

VK108, burial 4A/26 in Ljungbacka, Malmö, Sweden (Helene Wilhelmson, Yvonne Magnusson)

The site and excavation context: The Ljungbacka gravefield, in Lockarp near Malmö in Sweden, has been excavated multiple times since 1976. It contains graves from the Bronze Age and Iron Age, both cremations and inhumations, as was common practice in this region during those era. The grave field is described in multiple archive reports but only one details the grave in question (*121*). The archaeologist leading the excavations in Ljungbacka over the course of almost 30 years was Bengt-Åke Samuelsson who also published a synthesis and summary after the most recent excavation (*122*).

As a part of this specific project, the osteological and archaeological aspects of the find have been reanalysed and the results are presented here for the first time. Some of the methods used today are vastly different from the more traditional ones in use in 1976. The relevant aspects of the context should therefore be comparable to the other more recently analysed burials discussed in this supplement.

The grave 4/26, multiple and separate burials: This grave, featuring multiple single burials/features, was excavated in 1976. It was initially named 'Grav 4' (*121*) and then later 'Grav 26' (*122*). The descriptions in Table S17 are based on the information given in the original archive excavation report.

It appears that separate features, with separate human remains were collected as part of just one grave number. It is actually three different bone depositions (sometimes denoted A, B, and C; or just A and C and mixed, in the archive report and notes and bone boxes). These depositions further comprise more than three individuals. Two features are confirmed primary (articulated) burials from a taphonomic perspective: 4A (VK108) and 4C (VK109). These do not appear to have been in contact, judging from excavation plans and photographs. Another disturbed burial/bone assemblage was located, apparently on the same level as the upper of the complete burials. Due to the confusion in the different records of the grave (excavation plans, photos, report, notes and osteological notes; and notations on museum storage boxes), the graves are here denoted as on the boxes and as PS (primary burial, superior), PI (primary burial, inferior), and D (disturbed burial). The exact stratigraphical relation between the superior and inferior primary burials is highly interesting considering this context, but the existing documentation is not complete enough to allow such conclusions. This may very well be due to the fact that it was considered to be obviously two separate features during excavation, so no statement was made.

The skeletal remains: Individual 26A/4A/PS i.e., primary superior/upper burial (VK108, included in this study). The skeletal remains found in the boxes labelled 4A could be established to be those denoted 4A/26A in archive reports. The lower extremities and parts of the skull were somewhat eroded and more porous than the torso and pelvis. The lower extremities were fragmented and the feet appear to be missing. Most regions of the skeleton were observable and correlated with the documentation in excavation plans and photos (Fig. S27). With one exception, a lower second incisor, all skeletal material appears to be from a single individual. This tooth is most likely from a younger individual as it does not fit the empty alveoli for tooth 32 (FDI), has considerably less attrition than the other relevant teeth, and is of generally slightly larger size. This tooth is likely from a younger individual, a disturbed earlier burial for example.

This tooth has a very simple root and easily falls out of the alveoli once the soft tissue has decomposed. Many teeth were missing post mortem (FDI 11, 12, 14, 16, 17, 24, 26, 27, 32, 33, 34, 35, 36).

Using the criteria of sex estimation as summarised in Buikstra & Ubelaker (1994) (*123*), and most widely used today, this individual is scored in total as 'possible female' while showing mostly clearly feminine traits but also some intermediate ones (score 3). The sex, according to skull traits, should also be considered in the light of age, older females for example tending to score more towards the male spectrum when reaching menopause and beyond; i.e., age 55 and over (*124*), which is a relevant consideration in this case. This, the individual being over 55 years of age (see details below), could explain why some cranial traits score as possible female or indeterminate, and some as female. Further, as the pubis, which is considered most accurate in sex estimation in general (*125–127*) is missing, there is room for some doubt. The greater sciatic notch, which is available, is today considered to be variable (*128*, *129*) and scores both as clearly female and indeterminate for left and right side respectively. The concluding sex estimation is therefore given as 'possible female' rather than 'female'.

The age of the individual is estimated to be 55–75 years. This is based on the auricular surface of the ilium (the method of Buckberry and Chamberlain 2002 (*130*); score 15 for left side). This score gives an age of 55–79 years. This method has been revised in several studies (*131–135*) and is generally considered reliable but has a tendency to slightly over-score on older individuals. The age of the first rib was also estimated as 50–75 years using the methodology of Kunos *et al.* (1999) (*136*). Both these results were weighted together to an overall estimation of 55–75 years. Other general criteria, such as the bones being porous and light, likely resulting from osteoporosis onset, also concur with an older age. Further, periodontitis with secondary extensive root caries lesions and confirmed *in vitram* tooth loss, along with heavy attrition, also support this age estimate.

The fifth lumbar vertebrae was fused to the sacrum in an incomplete / partial sacralization of the vertebrae. Erosion and reactive bone in the anterior portion of L5-S1, could possibly be due to inflammation resulting from the sacralisation. There were no apparent skeletal manifestations of smallpox (as defined in the literature (*120*)) in this individual. This would however not be possible to find if the individual contracted the virus in adulthood, only if the infection was acquired in childhood (*120*).

In an earlier report (*137*), a potential sharp force trauma was noted just above the left eye in the site of the eyebrow. It was later described as likely lethal (*122*), despite the original report noting it as possibly caused by excavation damage. Today the margins appear irregular and unlike the very sharp margins of a sharp force trauma (Fig. S27). Further it would have been a very thin blade if the margins were sharp and now would have eroded to the more irregular current outline. It is more likely this is cracking of the bone as a result of taphonomy. The presence of sharp force trauma cannot be confirmed and as the original report also lists excavation damage as a likely cause, the current interpretation is that it is not a trauma.

Individual 4C/PI (primary burial, inferior/lower): The skeletal remains labelled 4C in the museum boxes appear to perfectly match the individual named 4B in drawings and the archive report. It is clearly not possible that these could be the mixed remains found at the same level as 4A, but is in fact a primary inferior/lower burial.

This individual had only clearly feminine traits of the skull and pelvis according to criteria in Buikstra & Ubelaker (1994) (*123*). The pubis was present and clearly female on all criteria. The age estimation is summarised as 60–72 years of age. The auricular surface (Buckberry &

Chamberlain (2002) (*130*); score 17 in left side) was determined to correspond to an age of 60–85 years. This method has been seen to over-score older ages (see discussion above). The pubic symphysis, albeit fragmented (*107*, *138*), was clearly from a late age phase (VI), giving an age of 48–72 years. The first rib was aged, according to Kunos *et al.* (1999) (*136*) to an age of 60+ years. The caput had eburnation and extensive remodelling (i.e., osteoarthritis) which would give an age of 76+ years. This is considered a potential pathology bias as only the left side was present and eburnation is not a criteria observed by Kunos *et al.* in their definition (*136*). The older age of this individual is supported by the very lightweight bones, probably resulting from osteoporosis. The most extensive periodontitis and *in vitram* tooth loss also correspond to an older individual. However, the pattern of tooth loss is complex. Both anterior and posterior teeth had been lost and the alveoli closed. In the maxilla mainly anterior teeth were missing (11-15; 21-23 FDI) and in the mandible (teeth 37, 36, 35, 47, 46, 45, 41, 42, possibly also 33, and 44). Such extensive tooth loss might be due to something other than simply advanced periodontitis, so the possibility of trauma should be considered. Further, there is a wide groove in superior-inferior direction on the lateral side of tooth 26 (FDI). This could possibly indicate a manual extraction of tooth 27 using tools. The groove floor is smooth, as if polished by tooth use for some time before death. Tooth 27 was clearly lost *in vitram* and the alveoli has healed. This individual had spondylolysis, with a fully separate arch, in the fifth lumbar vertebrae.

Grave 4 - mixed remains: The bones found in boxes labelled ‘grav 4 övre omrörda lagret’ (mixed remains) are likely those referred to being found in the vicinity of individual 4A/PS, aka VK108. These are the remains of at least two individuals, possibly three. No remains were possible to sex or age estimate, but they are clearly are adults of different sizes. The boxes also contained animal bones (fragment of diaphysis, medium sized mammal).

One finger phalanx (Ph I) was of a completely different preservation than all other bones from grave 4/26. It was very well preserved and had an oily appearance. This could be the ‘third’ individual in this feature with mixed remains.

The two or more adult individuals are represented by various skeletal elements. Most are direct duplicates of 4A/PS and 4C/PI and as such clearly represent different individuals. There will be no further detail here on these remains, other than taphonomical aspects.

Fragments of most long bones were possible to refit. Fragmentation is partly new (from excavation, storage, and handling) but to a large extent the fragments have eroded edges i.e., occurring before excavation. The bone surfaces are also eroded, but not in an apparent pattern, suggesting that movement occurred. This could be consistent with the grave soil of an earlier burial being disturbed and then redeposited.

These taphonomical aspects demonstrate that although pre-excavation fragmentation is extensive, commingling of parts of many individuals is not apparent. There are probably just three individuals represented here. The occurrence of different individuals in skull/teeth and of right and left hands of very different in sizes indicates two things:

1) human remains have been removed, but probably while articulated at least partly (very loose teeth still in jaw, more than one of the small bones from the same hand), and

2) one of the individuals was probably more intact than the other(s). Both upper and lower extremities, vertebral column, pelvis, hand, and foot likely come from the same individual as it is mainly the left side, with no apparent duplication.

DNA: The remains of the two primary burials, 4A and 4C, were included in a DNA-oriented study of Viking age human remains. This study concluded 4A/PS (VK108) and 4C/PI (VK109) were both genetically female (*99*).

A note on the burial form of VK108 (4A/PS): Prone burials are a phenomenon not uncommon in the Viking Age in Sweden, and also occur both earlier and later. There is a review of these burials in (*139*), arguing that they are likely related to lower social status. Recently, Toplak (*140*) refuted this interpretation as they are more common in some places, such as in the Viking Age on the island of Gotland in the Baltic Sea. Samuelsson (*122*) argues that burial 4A/26 was a murdered slave accompanying the lower burial i.e., 4C. He based this on the suggestion of a possible sharp force trauma to the skull (which cannot be confirmed today; see details above) and because the head was oriented to the south. Further, as the buried individual had a key, it is generally less likely it was a slave. It could either be the lady of the house (*141*) or possibly an especially trusted slave (see discussion in (*142*)). But in a burial context, keys are usually considered as indicating high social status (*141*). The social status is therefore considered here to be more difficult to interpret than the cases presented by Samuelson (*122*).

VK138: Galgedil, OBM 4520, grave AQQ skeleton x1572 (Marie Louise Schjellerup Jørkov)

Galgedil (Otterup Sogn, Lunde Herred, Odense Amt (Sb8)) is a Viking age cemetery (~800–1050 CE) located in the northern part of the Danish island of Funen. It was excavated by Odense Bys Museer in 1999-2005. The cemetery is situated on the top and down the southern and western slope of a small hill in a rolling, moraine landscape 5 km from the former seacoast. It revealed 54 graves containing 59 inhumations and 2 cremation burials of both males and females, and unsexed subadults of various ages. The cemetery was in use for ~250 years. There is no known nearby Viking age settlement at Galgedil, but it has been suggested that there must have been a sizable nearby settlement or that the cemetery was used by several smaller settlements.

The first phase, and also the largest group of burials, took place during the 9th and 10th century, and a phase of burials date to the 11th century. In the first phase of the cemetery, a minor group of six graves appear unusual, with wooden coffins, extraordinary artifacts, and deeper graves than other burials in the cemetery. Grave AQQ was one of these six graves. It contained two burials. The deepest placed burial contained the remains of an old adult male, 45+ years (x1773) buried with two knives, a whetstone and a belt with buckle and bracket with ornamentation inspired by the acanthus plant. VK138, a 25–30-year-old male (x1572) was buried 25 cm above this level, parallel to the other coffin, possibly on a wooden stretcher. He only had a knife with him. Based on the grave goods, the grave was dated to between 850 and 950 CE. Based on the differences in grave goods and positioning in the grave, it has been assumed that these were master and slave. The younger male showed a healed fracture on the left ulna and talus, and he had slight degenerative changes in his lumbar vertebrae. On the scapula, the acropion was separated from the rest of the bone, so-called ‘os acromiale’. This is either a congenital defect or may be related to mechanical stress.

Several individuals from Galgedil have been identified as non-locals based on strontium isotopes (*143*), however, both individuals from grave AQQ have been determined to be locals.

Previous scientific analysis of the remains has included light isotopes of carbon and nitrogen in collagen (10 samples) (*144*) and the radiocarbon determination of the age of 8 samples (*144*). In addition, mitochondrial aDNA was investigated in 10 samples from the cemetery (*145*).

Description of gene functions and mutations in the aVARV sequences

In the following text, numbers correspond to the final offset of the gene in CPXV-Gri/GER (*6*) and the name of the VACV-COP homolog (*6*), if applicable. If the gene is absent in VACV-COP, the CPXV naming convention is used.

Gene status is organized into four logical groups, denoted A-D, as shown in Fig. S17. Figure 2 shows the gene status for VARV, VARV-VD21, and aVARVs and the inferred gene status for internal phylogenetic tree nodes. A gene-inactivating mutation is described as 'novel' below if it does not occur in any of the 17 modern orthopoxvirus reference sequences considered here (Dataset 3).

Group A: Genes inactivated or absent in VARV, VARV-VD21, and the aVARV sequences.

This group contains a total of 38 genes, split into three sub-groups.

Group A1: Genes inactivated in VARV, VARV-VD21, and the aVARV sequences.

This group includes 15 genes: *B23R;B24R/C17L;C18L* (5392/218275), *20088*, *26765, C8L* (31175), *A25L* (150930), *A37R* (163060), *A39R* (165419), *A44L* (168244), *A53R* (176775), *B2R;B3R* (182986), *B12R* (191696), 193435*, *B16R* (194427), *C13L;C14L* (204801), and 215231.

B23R;B24R/C17L;C18L: Ankyrin (*6*). VARV-VD21, VARV, and all aVARV sequences have the same gene-inactivating mutation. The gene has a gene-inactivating mutation in TATV and is present in CMLV.

20088: CPXV019. Ankyrin (*6*). All aVARV sequences share a novel gene-inactivating mutation caused by a 1 base pair (bp) insertion. The gene-inactivating mutation in VARV-VD21 and VARV is identical, but is different from the aVARVs.

26765: CPXV025. ANK/F-box protein related to host range (*11*). VARV-VD21 has a 1 bp deletion, which leads to a 37 bp elongation of the sequence relative to VARV-SLN. The gene is of intermediate length between VARV-SLN and VARV-SAF. All aVARV sequences have a six bp insertion relative to VARV-SAF, which is also present in TATV. aVARV-VK388 and aVARV-VK281 share a stop codon with VARV-SAF. aVARV-VK382 has a 1 bp deletion relative to VARV-SAF, leading to a novel stop codon. If the 1 bp deletion is removed, aVARV-VK382 has the same stop codon as VARV-SAF. aVARV-VK470 has the same gene-inactivating mutation as VARV-SAF, but does not have coverage in the region of the 1 bp deletion in aVARV-VK382.

C8L: Unknown function (*6*). A novel, identical gene-inactivating mutation in VARV-VD21, and all aVARV sequences. The gene is absent in VARV-SAF, -SLN, and -KUW, while VARV-BRZ has a gene-inactivating mutation different from VARV-VD21. The gene has a gene-inactivating mutation in TATV and is present in CMLV

A25L: A-type inclusion protein (*6*). VARV-VD21 and all aVARV sequences share their gene-inactivating mutations with modern VARV,CMLV, and TATV.

A37R: Unknown (*6*). Modern VARV, VARV-VD21, and all aVARV sequences share the same gene-inactivating mutation. VARV-VD21 has an additional stop codon 3 amino acids upstream of this, caused by a single nucleotide substitution. The gene is functional in CMLV and TATV.

A39R: Semaphorin, secreted glycoprotein. Immunomodulator. Loss of the gene from VACV-COP does not affect growth *in vitro*, or virulence in a mouse intranasal model, and has

only a slight effect on lesion size in an intradermal mouse model. Expression of VACV-COP *A39* by VACV-WR, which does not naturally express the gene, leads to an increase in the severity and persistence of skin lesions after intradermal infection of mice (*146*). Histological examination of mouse skin suggests that *A39* has direct or indirect pro-inflammatory properties (*146*). aVARV-VK382 and aVARV-VK388 both share a 1 bp deletion relative to GPXV-Gri, which leads to a stop codon. However, aVARV-VK382 also has a 1 bp deletion upstream of the initial deletion, leading to an earlier stop codon. aVARV-VK281 and aVARV-VK470 have an identical novel stop codon. The stop codon in VARV-VD21 is also found in mVARV. CMLV and TATV also have different gene-inactivating mutations, differing both from each other and from mVARV and aVARV.

A44L: 3-beta-hydroxysteroid dehydrogenase (*147*). Deletion of *A44L* in VACV-WR does not affect viral replication in CV-1 cells *in vitro*, but the virus is attenuated in a mouse intranasal model (*147*). *A44* synthesizes steroid hormones, induces immunosuppression, and contributes to the virulence of VACV-WR in mice (*148*, *149*). This gene is inactivated in all sequenced VARV hitherto. VARV-VD21 has the same gene-inactivating mutation as VARV-SLN. All aVARV sequences share the same stop codon, but aVARV-VK388 and aVARV-VK470 have an additional stop codon upstream, due to a G → T substitution.

A53R: Also called *CrmC*, is a tumor necrosis factor receptor homolog that plays an important role in antiviral response and inflammation (*11*). Expression of the CPXV or VACV-USSR *A53R* gene in VACV-WR leads to increased virulence in a mouse intranasal model (*150*). Deletion of *A53R* in Tiantian VACV does not lead to a change of lesion size upon intradermal infection of rabbits, but results in virus attenuation in the mouse intranasal and intracranial model (*151*). All aVARV sequences have a 1 bp deletion at nt position 296, leading to identical stop codons in aVARV-VK382 and aVARV-VK388. aVARV-VK281 and aVARV-VK470 have an additional 1 bp insertion at position 165, leading to a different stop codon. The gene is absent in VARV-VD21 and the modern VARV sequences.

B2R;B3R: Poxvirus immune nuclease (poxin) / Schlafen-like. Immunomodulator (*152*). aVARV-VK281, aVARV-VK382, aVARV-VK388, and aVARV-VK470 share the same stop codon. VARV-VD21 contains a stop codon different to aVARVs but similar to stop codons in modern VARVs.

B12R: Has 33% amino acid similarity to *B1*, a serine/threonine kinase. However, *B12* lacks enzymatic activity (*153*). aVARV-VK281, aVARV-VK382, aVARV-VK388, aVARV-VK470, and VARV-VD21 have the same 1 bp insertion at position 239 of modern VARV-BRZ, -SLN, -SAF, and -KUW. However, the aVARV sequences all have an additional 5 bp deletion upstream of the insertion, leading to an earlier stop codon. The N-terminal end of aVARV-VK470 does not match the VARV-BRZ, -SLN, -SAF, and -KUW sequences.

*193435**: pCPXV0030. Surface glycoprotein (*6*). aVARV-VK382 and aVARV-388 have the same stop codon, but an additional stop codon in aVARV-VK388 terminates the gene earlier. The gene-inactivating mutations in VARV-VD21 and modern VARV are different. There is no read coverage in aVARV-VK281 or aVARV-VK470.

B16R: IL-1*β* receptor (*6*). Inactive in VACV-COP, but active in VACV-WR, where it is referred to as *B15R*. *"Deletion of B15R from vaccinia virus accelerated the appearance of symptoms of illness and mortality in intranasally infected mice, suggesting that the blockade of IL-1β by vaccinia virus can diminish the systemic acute phase response to infection and modulate the severity of the disease." (154)*. mVARV, VARV-VD21, and the aVARV sequences share the same gene-inactivating mutation.

C13L;C14L: Unknown (*6*). aVARV-VK281 and aVARV-VK470 share the same gene-inactivating mutation. The gene-inactivating mutations in the other aVARV sequences are different from each other. aVARV-VK382 and VARV-VD21 have the same gene-inactivating mutation as VARV.

215231: pCPXV0002. N-methly D-aspartate receptor-like protein (*6*). Gene-inactivating mutations agree in aVARV-VK382 and aVARV-VK388. The same 1 bp insertion is present in aVARV-VK281 and aVARV-VK470, but an earlier 1 bp deletion causes a different gene-inactivating mutation upstream to that in aVARV-VK382 and aVARV-388. The gene is absent in VARV-VD21 and modern VARVs.

Group A2: Genes inactivated in VARV, VARV-VD21, and certain or uncertain inactivations in some or all aVARV sequences.

This group includes four genes: *C9L* (29281), *A9L* (135109), *A52R* (175891), and 202711.

C9L: ANK/F-box protein related to host range (*11*), and an antagonist of the type I interferon response (*155*). VARV-VD21 has a novel stop codon relative to the modern VARV sequences. All aVARV sequences have a two bp insertion relative to TATV, leading to a novel stop codon. The gene is present in TATV and has a stop codon in CMLV.

A9L: Membrane protein (*6*). VARV-VD21 has an identical gene-inactivating mutation as VARV-BRZ. aVARV-VK388 and aVARV-VK382 have the same 27 bp insertion relative to VARV-BRZ and it is uncertain whether the gene is functional in these viruses. aVARV-VK281 and aVARV-VK470 have a novel gene-inactivating mutation. The sequences denoted as inactivated in Hendrickson et al., 2010 have lengths between 285 and 288 bp, whereas the functional sequences are between 300 and 420 bp long. aVARV-VK470 is 249 bp, and aVARV-VK382 and aVARV-VK388 are 315 bp long.

A52R: Intracellular Bcl2-like protein, blocks NFκB activation, and contributes to virulence in VACV (*156–158*). Deletion of *A52R* from VACV leads to attenuation in a murine intranasal model (*156*). The stop codon is identical in all aVARV sequences, but is marked as uncertain in aVARV-VK470, due to coverage of less than 3 reads.

202711: CPXV215. Kelch-like protein (*6*). VARV-VD21 and VARVs share the same gene-inactivating mutation. aVARV-VK281, aVARV-VK382, and aVARV-VK470 have the same gene-inactivating mutation. The stop codon in aVARV-VK281 is only supported by two reads, hence it is uncertain. aVARV-VK388 has the same gene-inactivating mutation as VARV-SLN, but it is caused by a 1 bp insertion at position 190 (4 reads) and a 1 bp deletion at position 554 (2 reads).

Group A3: Genes absent in VARV, VARV-VD21 and with no coverage in the aVARV sequences.

This group includes 19 genes: *8281*, *8318**, *9217*, *10160*, *12346*, *12699*, *14342*, *14947*, *15374*, *18018*, *19440*, *K4L* (43726), *K6L;K5L* (45027), *B9R* (188823), *B10R* (190433), *B11R* (190778), *202819*, *C15L/B21R* (212883), and *213858*.

Group B: Genes active in some or all aVARV sequences, but not in mVARV.

This group contains a total of 14 genes, split into two sub-groups.

Group B1: Genes inactivated in VARV and VARV-VD21, but present in all aVARV sequences.

This includes five genes: *C5L* (33086), *E7R* (68913), *B6R* (186591), *B7R* (187178), and *CrmE* (214439).

C5L: Genome uncoating and DNA replication factor (*41*). Coverage of reference sequence (CPXV-Gri): aVARV-VK281: 99.4%, aVARV-VK382: 100%, aVARV-VK388: 100%, aVARV-VK470: 100%. The gene-inactivating mutation in VARV and VARV-VD21 is identical.

E7R: EEV myristylated soluble protein (*159*). Inactivation does not affect known biological properties of the virus (*160*). Coverage of reference sequence (TATV): aVARV-VK281: 98.6%, aVARV-VK382: 100%, aVARV-VK388: 100%, aVARV-VK470: 100%. The gene-inactivating mutation in VARV and VARV-VD21 is identical.

B6R: Ankyrin (*6*). Coverage of reference sequence (TATV): aVARV-VK281: 100%, aVARV-VK382: 100%, aVARV-VK388: 100%, aVARV-VK470: 100%. The gene-inactivating mutation in VARV and VARV-VD21 is identical.

B7R: Endoplasmic reticulum-associated, putative chemokine binding protein. A virulence factor in VACV-WR, its deletion leads to the attenuation (smaller lesions) of VACV-WR in a murine intradermal model, but not in the murine intranasal model (*161*). Coverage of reference sequence (TATV): aVARV-VK281: 99.5%, aVARV-VK382: 100%, aVARV-VK388: 100%, aVARV-VK470: 99% (the first two amino acids do not have coverage). The gene-inactivating mutation in VARV and VARV-VD21 is identical.

CrmE: pCPXV0002. Tumor necrosis factor (TNF) receptor homolog. *CrmE* inhibits the cytotoxic and apoptotic activities of human, but not mouse or rat, TNF *in vitro*. In a murine intranasal model, VACV-USSR recombinants lacking *CrmE* are attenuated. Expression of *CrmE* in VACV-WR enhances virulence in the murine model (*150*). Absent in both VARV-VD21 and modern VARV. The gene appears to be functional in aVARV-VK382, aVARV-VK388, aVARV-VK281, and aVARV-VK470. *CrmE* sequences present in the aVARVs are most closely related to CPXV-Gri and CPXV-Ger. Coverage of reference sequence (CPXV-Ger): aVARV-VK281: 96.2%, aVARV-VK382: 100%, aVARV-VK388: 100%, aVARV-VK470: 100%. Active *CrmE* orthologs are only found in CPXV-Gri, CPXV-Ger, VACV-LIS, VACV-Lc16m8, VACV-Lc16mO, and VACV-USSR strains, but are absent in all other VACV strains as well as in HSPV, RPXV, VARV, and TATV, and are truncated in CMLV and ECTV (*11*). Gene-inactivating mutations are present in identical positions in MPXV, ECTV, and CPXV-BR, suggesting that the inactivation arose either in an ancestral virus, or as a result of horizontal gene transfer or recombination events (*11*), in which case aVARV-VK388 and aVARV-VK470 may have also received their version of *CrmE* in a similar fashion.

Group B2: Genes inactivated in VARV and VARV-VD21, but present in some but not all aVARV sequences.

This group includes nine genes: 22682, *C2L* (35634), *K1L* (41155), *F3L* (47695), *A35R* (161464), *A40R* (165942), *168145, A55R* (178726), and *A57R* (180331).

22682: CPXV020. Unknown function (*6*). aVARV-VK382 and aVARV-VK388 share the same gene-inactivating mutation. VARV-VD21 and VARV-SLN also have an identical stop codon. aVARV-VK281 and aVARV-VK470 are four amino acids shorter than CPXV-BR, and may still be functional.

C2L: Kelch-like proteins with putative host-range function (*162*). Deletion of *C2L* in VACV has no effect on growth *in vitro*, but has altered plaque morphology (*31*). Deletion of *C2L*

in VACV also does not affect virulence in a mouse intranasal model, but leads to larger lesions and more cell infiltration in a mouse intradermal model (*31*). Thus, *C2L* contributes to cell cytopathic effect and reduces immunopathology *in vitro (31)*. aVARV-VK281 and aVARV-VK470 have an identical novel gene-inactivating mutation. Coverage of reference sequence (CMLV): aVARV-VK382: 100%, aVARV-VK388: 100%. VARV-VD21 and VARV have the same gene-inactivating mutation.

K1L: NFκB inhibitor, and human SAMD9 inhibitor (*37*, *163*). Plays a role in antiviral response and inflammation, and is a known host-range gene (*11*) since its deletion in VACV-COP results in a replication defect in rabbit RK13 cells (*39*). The presence of a functional *K1L* gene in aVARV-VK382 and aVARV-VK388, as well as the presence of different gene-inactivating mutations in CMLV, TATV, and VARV (*11*) suggest that the gene was lost independently in CMLV, TATV, and VARV after the divergence of the aVARV clade and mVARV ~1.7 kya. Coverage of reference sequence (CPXV-Gri): aVARV-VK382: 100%, aVARV-VK388: 100%. aVARV-VK281 and aVARV-VK470 have the same novel gene-inactivating mutation. VARV-VD21 and VARV have a different novel gene-inactivating mutation than aVARV-VK281 and -VK470.

F3L: Kelch-like protein (*6*). Deletion of *F3L* in VACV leads to smaller lesions in a mouse intradermal model, as indicated by smaller lesions, and alterations in the innate immune response (*33*). Coverage of reference sequence (CMLV): aVARV-VK382: 100%, aVARV-VK388: 100%. A novel stop codon is present in aVARV-VK281 and aVARV-VK470. VARV-VD21 and VARV have the same gene-inactivating mutation, different to aVARV-VK281 and aVARV-VK470.

A35R: Virulence factor. Inhibits MHC class II-restricted antigen presentation, immune priming of T lymphocytes, and subsequent chemokine and cytokine synthesis (*164*). Coverage of reference sequence (RPXV): aVARV-VK388: 100%, aVARV-VK470: 100%. Two novel stop codons are present in aVARV-VK281 and aVARV-VK382. VARV-VD21 and VARV have the same gene-inactivating mutation.

A40R: Surface glycoprotein, immune modulator, related to C-type lectin-like proteins. Disruption of the *A40R* gene in VACV did not affect virus plaque size, *in vitro* growth rate and titre, EEV formation, or virus virulence in a murine intranasal model, but led to smaller lesions in the mouse intradermal model (*165*, *166*). No coverage of aVARV-VK281 and aVARV-VK470. In aVARV-VK382, a 1 bp deletion at nt position 441 leads to an extension of the gene by eight amino acids relative to CPXV. VARV-VD21 and VARV-BRZ share an identical stop codon. aVARV-VK388 has a novel stop codon.

168145: CPXV181. Unknown function (*6*). Coverage of reference sequence (CMLV): aVARV-VK382: 100%, aVARV-VK388: 100%. VARV-VD21 and VARV have the same gene-inactivating mutation. The gene-inactivating mutation in aVARV-VK281 and aVARV-VK470 is identical to the one in VARV.

A55R: Kelch-like protein. *A55R* is a NFkB inhibitor (*167*). Deletion of *A55R* in VACV-WR does not affect growth rate *in vitro* but leads to a different plaque morphology and cytopathic effect. It also leads to larger lesions in a murine intradermal model, hence affecting the host response to infection *in vivo (32)*. Coverage of reference sequence (CPXV-Gri): aVARV-VK281: 97.9%, aVARV-VK382: 100%, aVARV-VK388: 100%. An uncertain novel stop codon in aVARV-VK470 is covered by two reads. VARV-VD21 and VARV have the same gene-inactivating mutation.

A57R: Guanylate kinase. An inactive gene in VACV-COP and VACV-WR, with the same gene-inactivating mutation in mVARVs (*13*). *A57* is predicted to only be functional if it has an additional 5' region that contains an ATP binding domain (*168*). In VARV, VARV-VD21, VACV-COP, VACV-WR, VACV-MVA, and RPXV, the gene is inactivated by an 11 bp deletion relative to CPXV. In HPXV, ECTV, TATV, CMLV, MPXV-ZAI, MPXV-WR, aVARV-VK281, and aVARV-VK470, the gene is inactivated by nucleotide substitutions leading to a stop codon, or 1 bp insertions or deletions that cause a frameshift. The mutation that inactivates MPXV-ZAI and MPXV-WR is identical, and the deletion that inactivates aVARV-VK281 and aVARV-VK470 is also present in ECTV, although the gene in ECTV is inactivated upstream due to a substitution. CPXV-Ger, CPXV-Gri, CPXV-BR, aVARV-VK382, and aVARV-VK388 all have a 5' region with the ATP binding domain. VARV-VD21 and VARV have the same gene-inactivating mutation. Coverage of reference sequence (CPXV-Ger): aVARV-VK382: 100%, aVARV-VK388: 100%.

Group C: Genes present in VARV and VARV-VD21, but absent in some aVARV sequences.

Group C1: Genes present in VARV and VARV-VD21, and inactivations in some but not all aVARV sequences.

This group includes three genes: *CrmB* (*B28R/C22,* 2286/221381), *C1L* (37242) and *E5R* (66488).

CrmB: Secreted decoy tumor necrosis factor receptor (N-terminal domain) and highly-specific binding protein for chemokines (C-terminal domain) (*6*, *35*). A novel stop codon is present in aVARV-VK281 and aVARV-VK470. In aVARV-VK388, the start codon is replaced by a valine. Coverage of reference sequence (VARV-SLN): aVARV-VD21: 100%, aVARV-VK382: 100%, aVARV-VK388: 100%, aVARV-VK470: 56.3%.

C1L: Unknown function. Bcl2-like (*6*). Coverage of reference sequence (VARV-KUW): aVARV-VK281: 100%, aVARV-VK382: 100%, aVARV-VK470: 100%, VARV-VD21: 100%. A novel stop codon is present in aVARV-VK388.

E5R: Virosome component protein (*6*, *169*). Functional in aVARV-VK388. Novel identical stop codons in aVARV-VK281, aVARV-VK382, and aVARV-VK470. Coverage of reference sequence (TATV): aVARV-VK388: 100%, VARV-VD21: 100%.

Group C2: Genes present in VARV and VARV-VD21, and uncertain inactivations in aVARV sequences.

This group includes four genes: *F10L* (52696), *I8R* (85206), *B20R* (200936), and *210863*.

F10L: Serine/Threonine kinase (*6*). The gene is present in all aVARV but has an uncertain inactivation in aVARV-VK281, due to a C → T mutation covered by one read. Coverage of reference sequence (VARV-SLN): aVARV-VD21: 100%, aVARV-VK382: 100%, aVARV-VK388: 100%, aVARV-VK470: 100%.

I8R: RNA helicase (*6*). The gene is present in all aVARV but has an uncertain inactivation in aVARV-VK281, due to a T → A mutation covered by one read. Coverage of reference sequence (CMLV): aVARV-VD21: 100%, aVARV-VK382: 100%, aVARV-VK388: 99.9%, aVARV-VK470: 100%.

B20R: Ankyrin (*6*). The gene is present in all aVARV but has an uncertain inactivation in aVARV-VK281, due to a 1 bp substitution supported by two reads. Coverage of reference

sequence (VARV-BRZ): aVARV-VD21: 100%, aVARV-VK382: 100%, aVARV-VK388: 100%, aVARV-VK470: 100%.

210863: CPXV219. Unknown function (*6*). The gene is present in all aVARV but has an uncertain inactivation in aVARV-VK281 due to a 1 bp insertion covered by one read. Coverage of reference sequence (TATV): aVARV-VD21: 100%, aVARV-VK382: 100%, aVARV-VK388: 99.6%, aVARV-VK470: 99%.

Group D: Genes present in VARV, VARV-VD21 and the aVARV sequences.

This group includes 155 genes: *C23L/B29R* (1393/222274), *C19;C20L/B25/B27R* (3420/220247), *C16L/B22R* (7652/216015), *C11R* (23772), *C10L* (23924), *26154*, *26324*, *28912*, *C7L* (31795), *32306*, *C6L* (32479), *C4L* (33774), *C3L* (34788), *N1L* (37924), *N2L* (38413), *M1L* (38982), *M2L* (40901), *K2L* (42228), *K3L* (43400), *K7R* (46445), *F1L* (46512), *F2L* (47228), *F4L* (49163), *F5L* (50154), *F6L* (51155), *F7L* (51395), *F8L* (51813), *F9L* (52071), *F11L* (54038), *F12L* (55145), *F13L* (57091), *F14L* (58227), *F14.5L* (58498), *F15L* (58720), *F16L* (59203), *E1L* (60263), *F17R* (60266), *E2L* (61699), *E3L* (64011), *E4L* (64637), *E6R* (68330), *E8R* (69862), *E9L* (69869), *E11L* (73203), *E10R* (73208), *O1L* (73579), *O2L* (75627), *75977*, *I1L* (76099), *I2L* (77044), *I3L* (77266), *I4L* (78158), *I5L* (80500), *I6L* (80758), *I7L* (81899), *G1L* (85210), *G3L* (86982), *G4L* (87943), *G2R* (87973), *G5R* (89624), *G5.5R* (89823), *G7L* (90287), *G6R* (90322), *G8R* (92215), *G9R* (93257), *L1R* (94010), *L3L* (94296), *L2R* (94320), *L4R* (96128), *L5R* (96524), *J1R* (96942), *J2R* (97491), *J3R* (98558), *J4R* (99030), *J5R* (99091), *H1L* (103456), *J6R* (103459), *H2R* (104554), *H3L* (104557), *H4L* (105532), *H5R* (108743), *H6R* (109688), *H7R* (110165), *D2L* (112702), *D1R* (112743), *D3R* (113848), *D4R* (114504), *D5R* (116893), *D6R* (118847), *D8L* (119322), *D7R* (119359), *D9R* (120919), *D10R* (121662), *D11L* (121663), *D12L* (123593), *D13L* (124487), *A1L* (126166), *A2L* (126639), *A2.5L* (127310), *A3L* (127555), *A4L* (129542), *A6L* (130922), *A5R* (130925), *A7L* (132064), *A8R* (135116), *A10R* (135433), *A11R* (139079), *A12L* (139081), *A13L* (139689), *A14L* (140005), *A14.5L* (140294), *A15L* (140445), *A16L* (140713), *A17L* (141849), *A19L* (143937), *A18R* (143956), *A21L* (144168), *A20R* (145800), *A22R* (146293), *A23R* (147461), *A24R* (150952), *A26L* (154815), *A27L* (156422), *A28L* (156755), *A29L* (157196), *A30L* (158076), *A30.5L* (158342), *A32L* (158873), *A31R* (158906), *A33R* (160360), *A34R* (160890), *A36R* (162202), *A38L* (163362), *163365*, *A41L* (166040), *A42R* (167273), *A43R* (167901), *A45R* (169708), *A46R* (170420), *A47L* (170515), *A48R* (171963), *A49R* (172501), *A50R* (174192), *A51R* (175249), *A56R* (179720), *B1R* (181384), *B4R* (184887), *B5R* (185944), *B8R* (188071), *B13R;B14R* (192830), *B15R* (193355), *B17L* (194473), *B18R* (197359), *B19R* (198481), and *C12L* (204069).

Modeling gene inactivation (or loss) over time

Figure 3 suggests that it might be appropriate to infer a linear rate of gene inactivation (or loss) over time following the virus host switch into humans. We have chosen not to do so because many of the following factors, about which nothing is currently known, would be inappropriately ignored (i.e., incorporated as implicit assumptions) under a linear interpretation or other simple model:

- That inactivation of unnecessary (in the new host) host-specific genes is independent.
- That the virus was only present in one host during the observed period of gene loss (cf MPXV).

- That the virus was in a broad-host-range host (with many active genes) immediately before the zoonosis.
- That selectively neutral (in the new host) genes are not inactivated more rapidly in the period soon after host switches.
- That zoonosis occurred only once, or that all viruses in the analysis are descendants of the same zoonotic event (and thus have evolved in the same host).
- That the MRCA of the observed viruses had a human host.
- That (if the model allows for other hosts) the rate of gene inactivation in those other hosts does not differ significantly from the rate in humans.
- That inactivated (via a stop codon) partial genes have no function.
- That mutations in non-coding regions of the genome had no influence on the impact of gene inactivation.

The linear appearance of the current data is driven by three sample clusters: the mVARV sequences and two groups of just two aVARV samples each (VK382 and VK388 in one and context-dated samples VK281 and VK470 in the other). Figure 3 makes it clear that the confidence interval around the x-axis (zero gene inactivations, not shown) would be very broad. Until additional data is available, it makes the most sense to be cautious regarding the overall progression of gene loss. There may have been a single zoonosis, and the rate of gene inactivation in humans may indeed have been linear, but other scenarios are quite plausible. That a simple linear model based on multiple unrealistic assumptions might fit the data does not provide additional insight and at worst may mislead. The true state of affairs regarding most of the above factors is likely to remain unknown and untestable in human hosts, given the pathogen involved.

Global smallpox mortality estimates

The literature on variola virus contains frequent references to an estimate that smallpox killed 300-500 million (M) people in the 20th century alone. Following citations, when given, we have tried to find the source of this estimate, without success. It may be that the estimate was originally an accepted informal calculation passed around by word of mouth, which gradually made its way into the literature. Fortunately, there are some presumably reliable numbers that lead to an estimate in this range. Versions of the WHO fact sheet on smallpox available at least as late as Oct 23, 2012 (*170*) give an estimated number of ~50 M new cases annually in the early 1950s, and state that this had dropped to 10-15 M cases a year by 1967. These figures are also given by Fenner, Henderson, et al. (*8*), who further state that there were an estimated 1.5-2 M deaths annually by 1967. If these 1967 numbers are correct, they represent a global case fatality rate of 10-20%. Applying this to the case number of ~50 M for the early 1950s gives an annual mortality figure in the region of 5-10 M at that time. If that is extrapolated to cover the first 70 years of the 20th century, a total fatality count of 350-700 M results. Given the assumptions present in that calculation, the oft-quoted estimate of 300-500 M does not appear unreasonable.

The same cannot be said for the occasional assertion that smallpox alone killed more people than all other infectious diseases combined. To pick just one other pathogen, Hepatitis B virus, even after the development of an effective and widely deployed vaccine, is still responsible for almost 900,000 deaths a year (*171*) and so can be expected to have been responsible for at least

100 M deaths by itself in the 20th century. Based on the written record (or the conspicuous absence of any mention), smallpox does not seem to have caused massive mortality before the 17th century (*172*), during which time epidemics of other infectious diseases were frequent, sometimes with estimated mortality rates far higher than smallpox at its worst (e.g., > 50% in some *Yersinia pestis* outbreaks). We thank the anonymous reviewer who prevented us from repeating this unsupported claim.

Figures

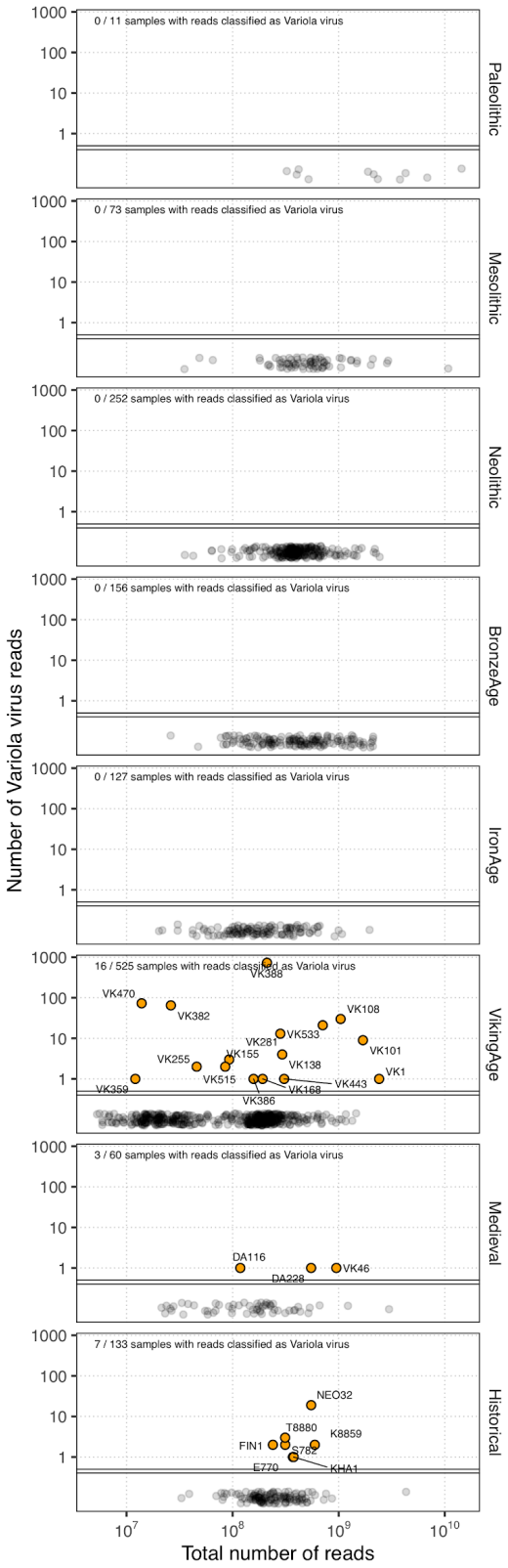

0 / 11 samples with reads classified as Variola virus
0 / 73 samples with reads classified as Variola virus
0 / 252 samples with reads classified as Variola virus
0 / 156 samples with reads classified as Variola virus
0 / 127 samples with reads classified as Variola virus
16 / 525 samples with reads classified as Variola virus
3 / 60 samples with reads classified as Variola virus
7 / 133 samples with reads classified as Variola virus
Paleolithic
Mesolithic
Neolithic
BronzeAge
IronAge
VikingAge
Medieval
Historical
Number of Variola virus reads
Total number of reads
1000
100
10
1
10^7
10^8
10^9
10^{10}
VK388
VK470
VK382
VK108
VK533
VK281
VK101
VK155
VK255
VK138
VK1
VK515
VK443
VK359
VK168
VK386
DA116
VK46
DA228
NEO32
T8880
K8859
FIN1
S782
E770
KHA1

Fig. S1. Kraken screening results. Panels show the number of reads classified as variola virus against the total number of reads sequenced for each screened sample, grouped by age period. Samples with evidence for VARV are indicated by orange circles and labelled by sample name. The distributions of total read count for samples with no VARV-matching reads are shown as grey point clouds below the double line (y-axis jitter added to aid visualization). The total sample count (1337) in the figure differs from the overall number of samples screened (1867) because only samples for which dating has thus far been clearly established are included above.

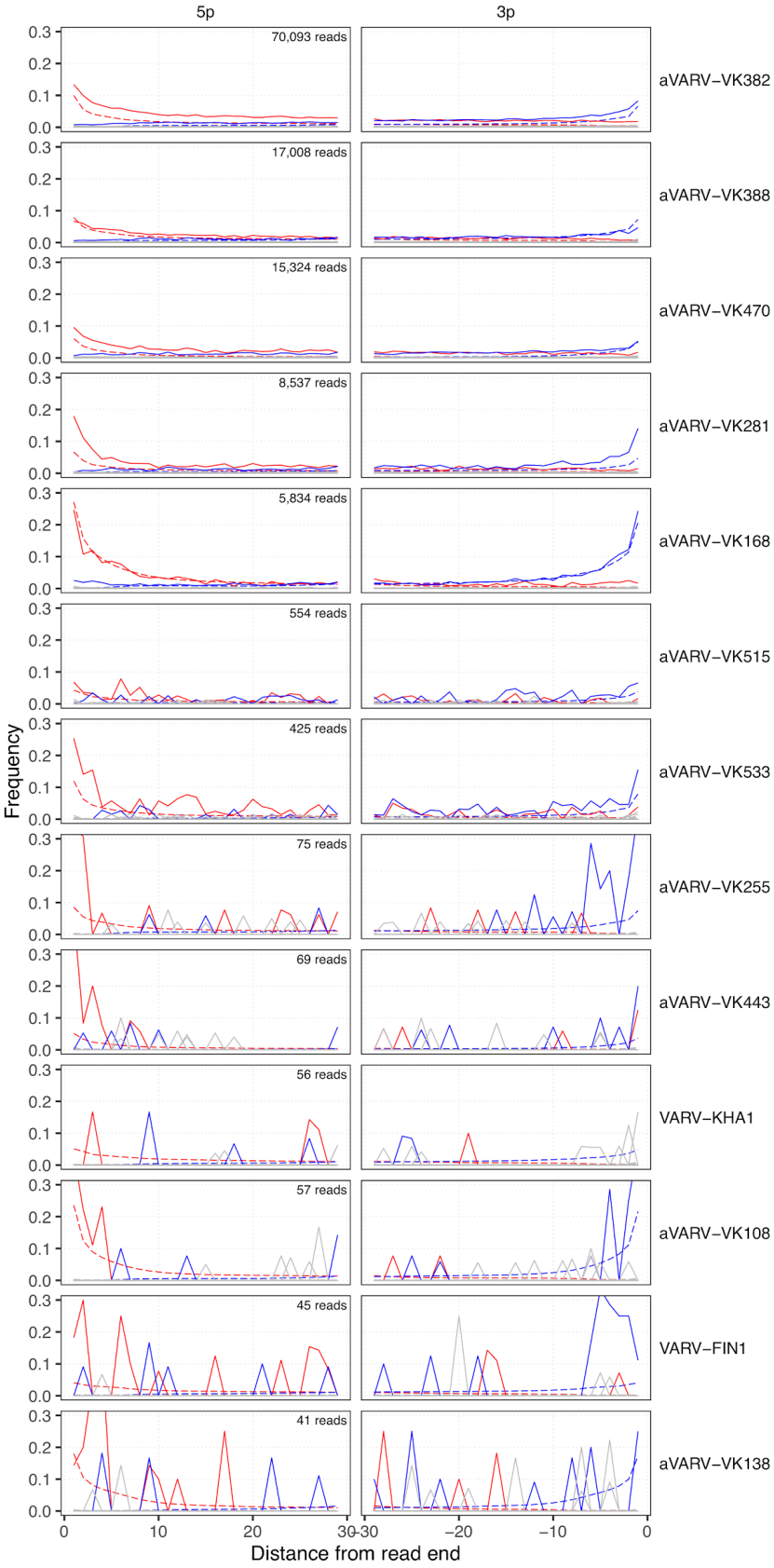

5p
3p
70,093 reads
17,008 reads
15,324 reads
8,537 reads
5,834 reads
554 reads
425 reads
75 reads
69 reads
56 reads
57 reads
45 reads
41 reads
aVARV−VK382
aVARV−VK388
aVARV−VK470
aVARV−VK281
aVARV−VK168
aVARV−VK515
aVARV−VK533
aVARV−VK255
aVARV−VK443
VARV−KHA1
aVARV−VK108
VARV−FIN1
aVARV−VK138
Frequency
0.0
0.1
0.2
0.3
0
10
20
30
−30
−20
−10
0
Distance from read end

Fig. S2. Ancient DNA damage patterns. The frequencies of the mismatches observed between the post-capture reads and the TATV reference sequence (GenBank accession no.: NC_008291.1) are shown as a function of distance from the 5' (5p) and 3' (3p) ends (solid lines). C → T (5') and G → A (3') mutations are shown in red and blue, respectively. All other possible mismatches are shown in grey. Mismatch rates for reads mapping to the human genome are shown (dashed lines) for comparison. Damage plots are made using mapDamage based on an --end-to-end very sensitive alignment made using Bowtie2. Read counts and characteristics are as described in the caption of Table S2.

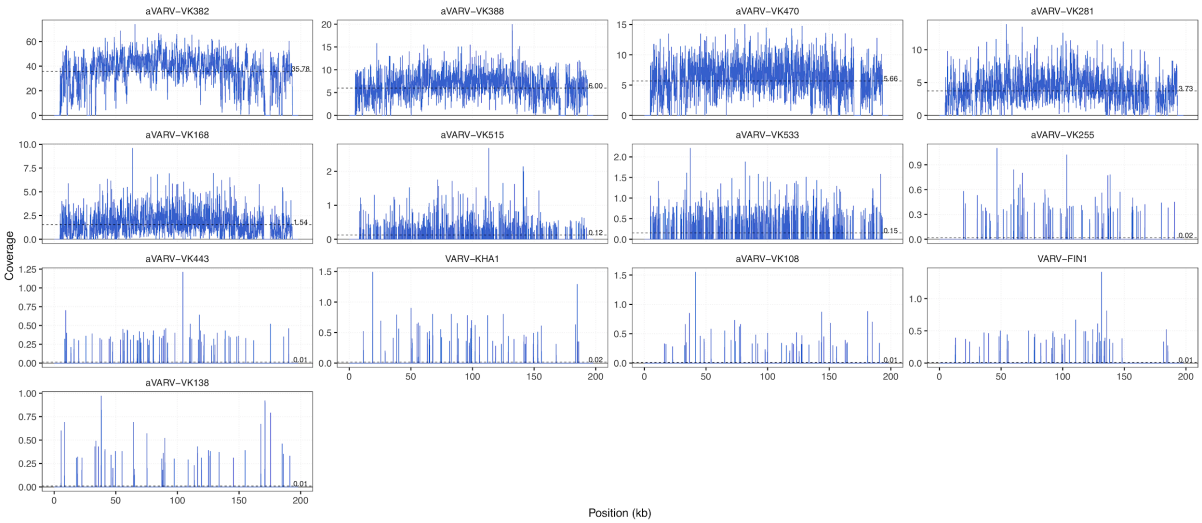


Fig. S3. Genomic coverage depth. Average genomic coverage depth in 1000 nt windows across the TATV reference sequence (GenBank accession no.: NC_008291.1), obtained using mosdepth, based on a Bowtie2 --end-to-end mapping. Dashed horizontal lines indicate genome-wide average coverage depth.

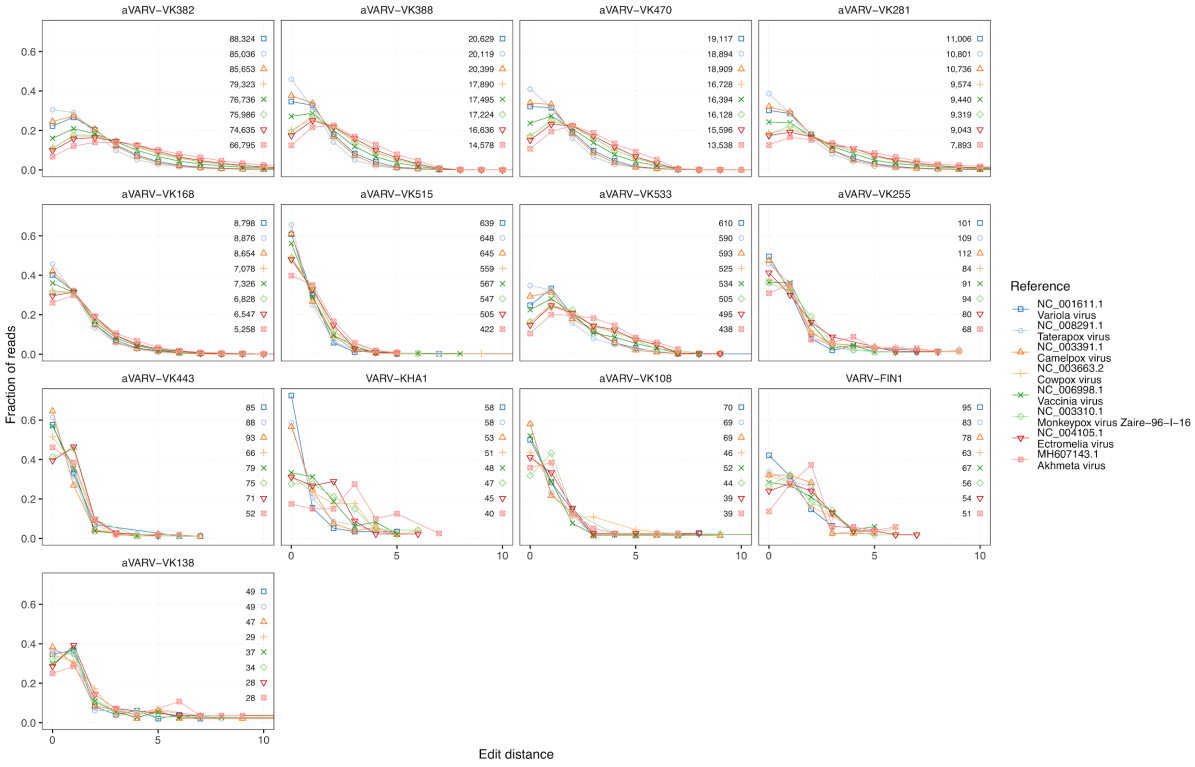
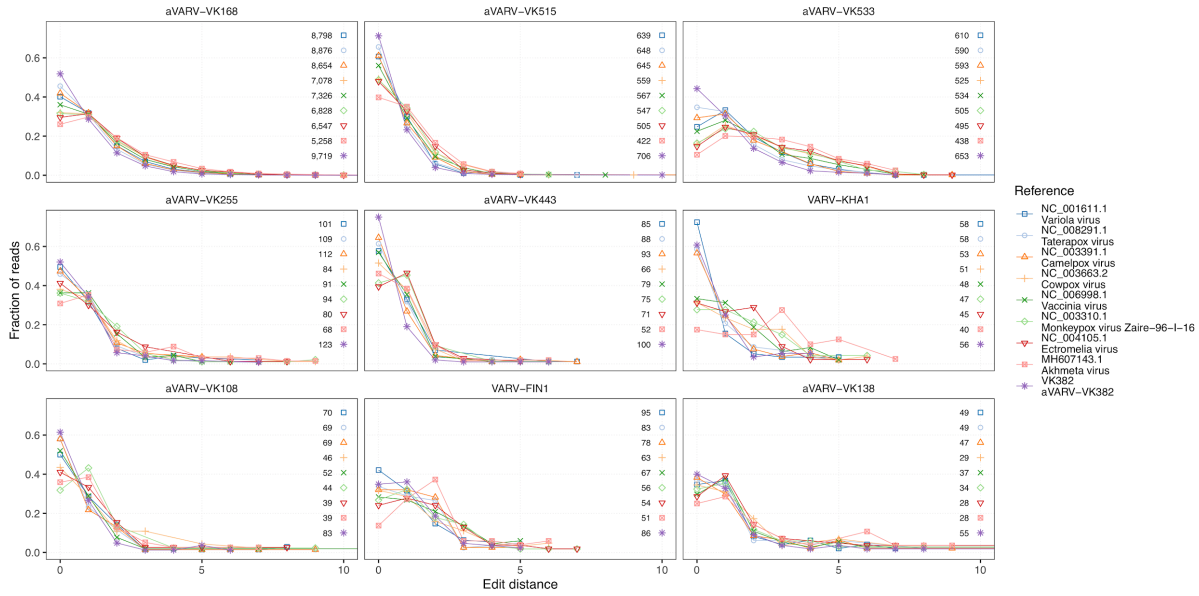


Fig. S4. Edit distance distributions. Panels show (post-capture) read edit distance distributions for each sample, mapped against eight different orthopoxvirus reference genomes (Bowtie2, local alignment mode). Numbers in each plot indicate the total number of reads for each sample and reference genome. Distributions are skewed furthest to the left (i.e., towards the lowest edit distance) for the most-closely related reference genome. Reads from Viking age samples have lowest edit distance to either TATV or CMLV, whereas the 19th century samples (FIN1 and KHA1) have lowest distance to VARV.

A



B

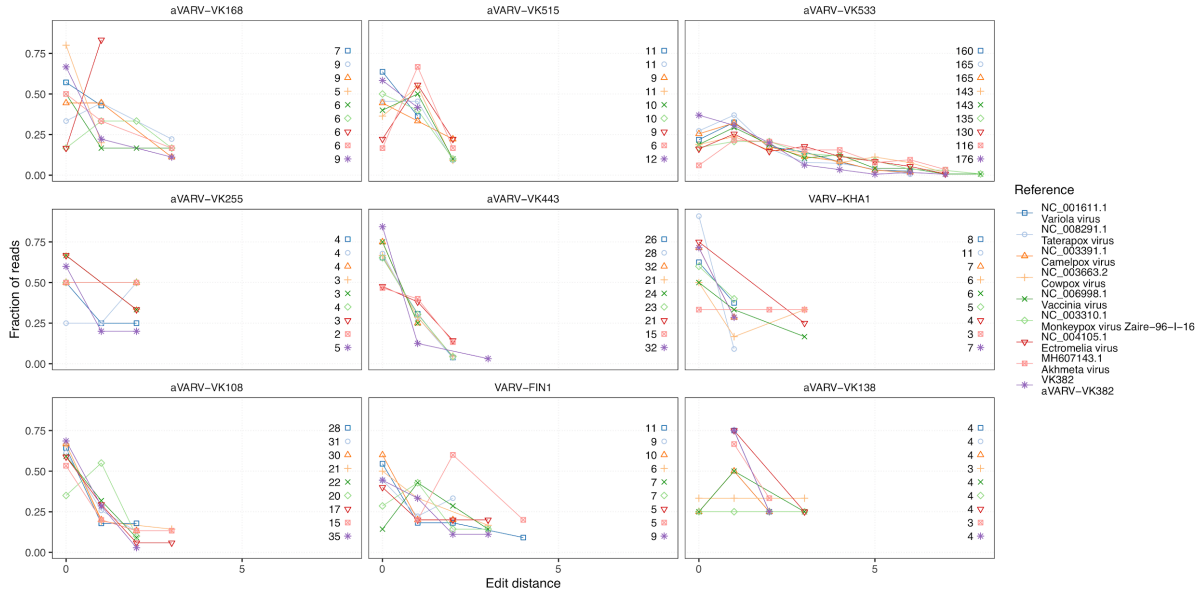


Fig. S5. Edit distances of reads from lower-coverage samples mapped against eight different orthopoxvirus reference sequences and the aVARV-VK382 consensus sequence. Numbers in each plot indicate the total number of reads for each sample and reference genome. **A) Edit distances of post-capture reads.** Lower-coverage samples from the Viking Age have lowest (post-capture) read edit distance to aVARV-VK382 consensus, and samples dated to the 19th century have lowest edit distance to VARV, providing read-level confirmation of the same finding at the overall consensus level (Figs. 1B, S13–S15). **B) Edit distances of pre-capture reads.** The majority of the samples show noisy patterns due to low numbers (10 or less) of pre-capture reads. However, the three samples with greater than 20 pre-capture reads (VK108, VK443, and particularly VK533), show the same pattern of lowest edit distance to the aVARV-VK382 consensus reference, consistent with the results from the merged pre- and post-capture data.

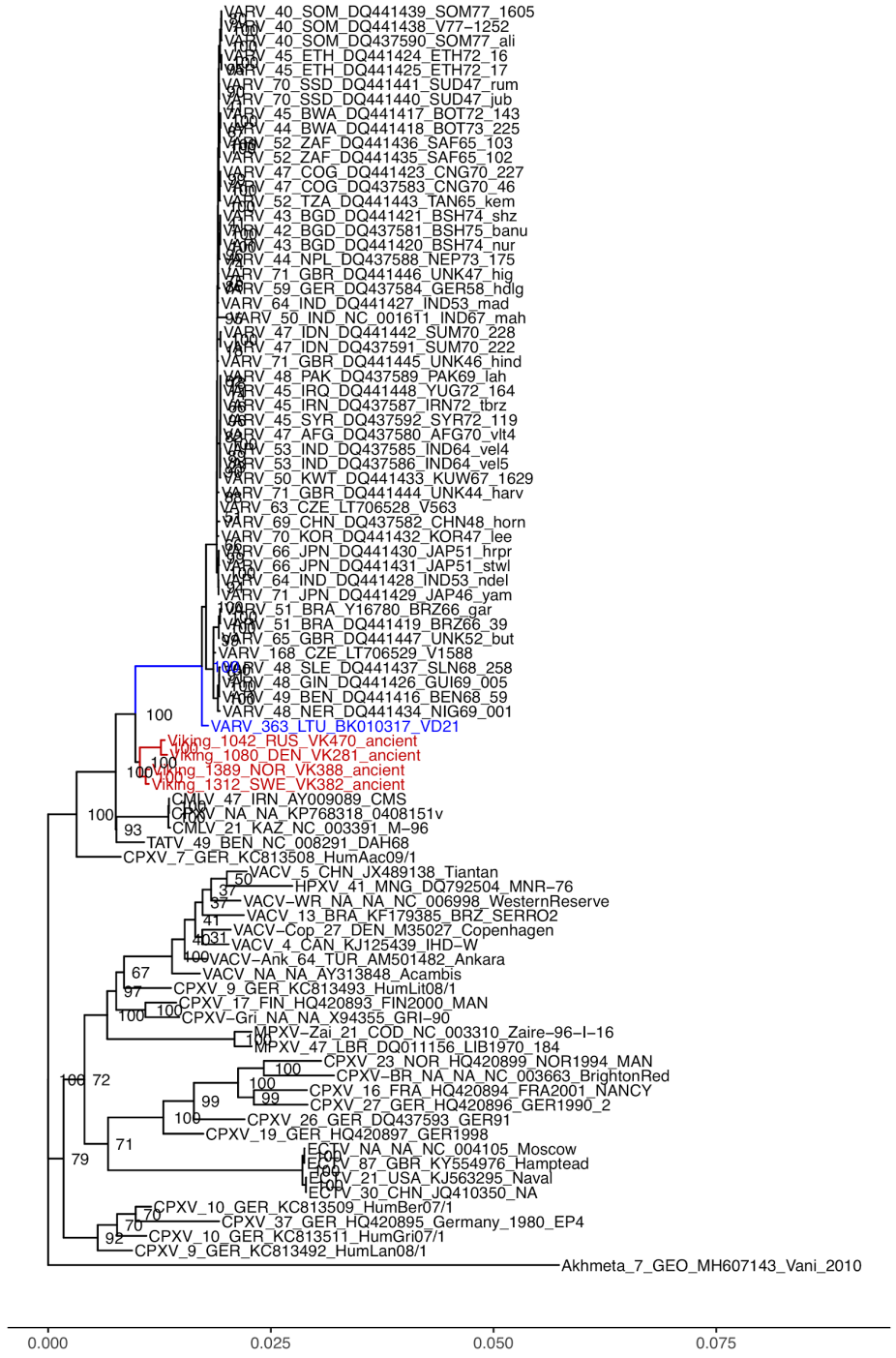


Fig. S6. Maximum likelihood tree of the orthopoxviruses and the Viking age higher-coverage samples used for EPA-ng analysis. Sequences were aligned using MAFFT. The tree was inferred using RAxML-NG with a K81uf+G+I substitution model, and an ML estimate of tree topology, branch lengths, substitution rates, and nucleotide frequencies, with support estimated using 1000 bootstrap replicates. The aVARV sequences are shown in red, the ancient

sequence from Duggan *et al.*, 2016 in blue (*16*). Taxon name fields indicate: virus/historical period, sample age relative to 2017, country abbreviation of sequence origin and region of sequence origin, as determined by the Standard country or area codes for statistical use, GenBank accession number/sample identifier, additional remarks. The x-axis denotes substitutions per site. Lower-coverage samples were later placed on this tree using EPA-ng, as described in Materials and Methods. The tree is rooted with Akhmeta_7_GEO_MH607143_Vani_2010.

A

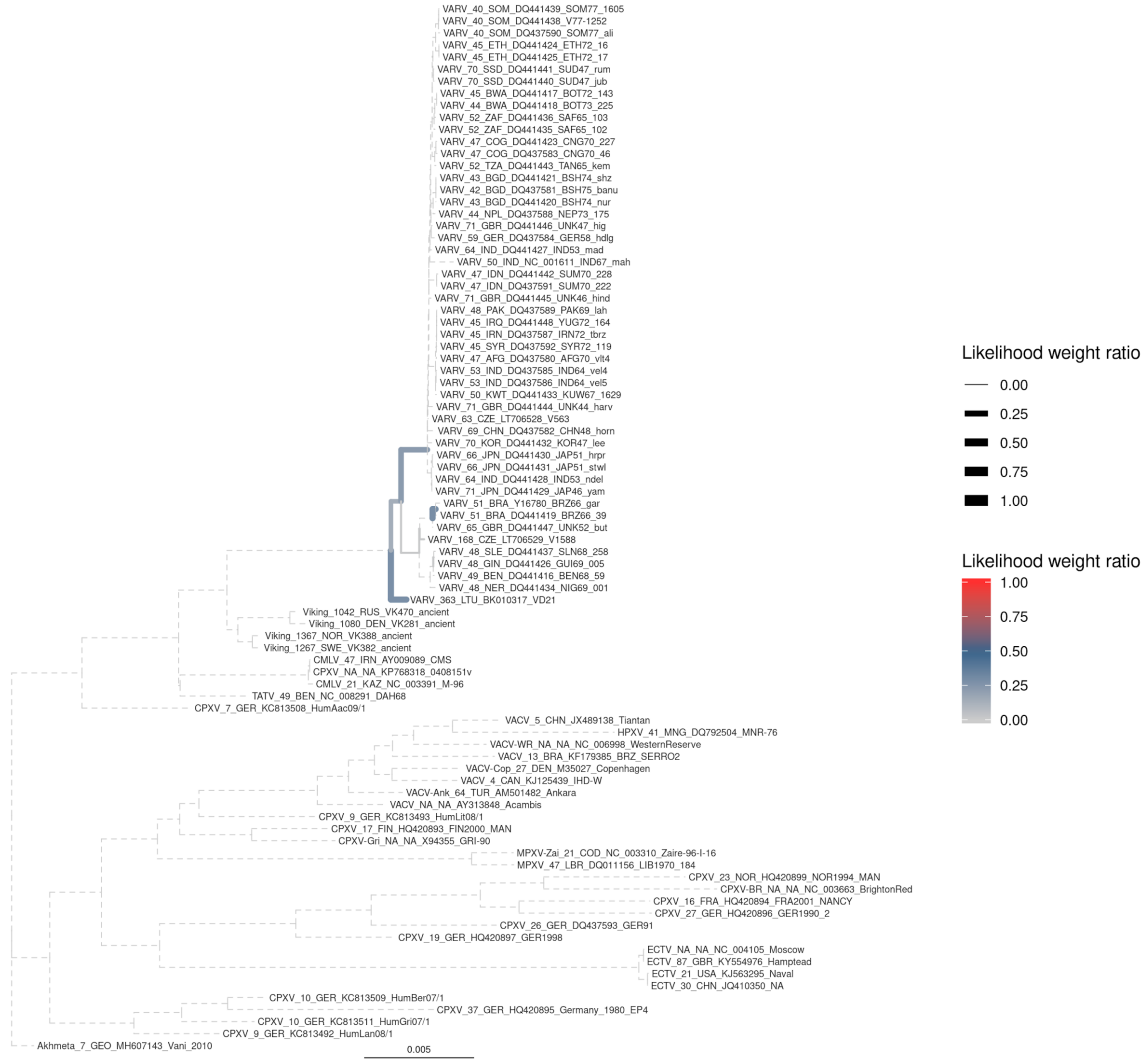
aVARV-VK168

B

aVARV-VK533

C

VARV-FIN1

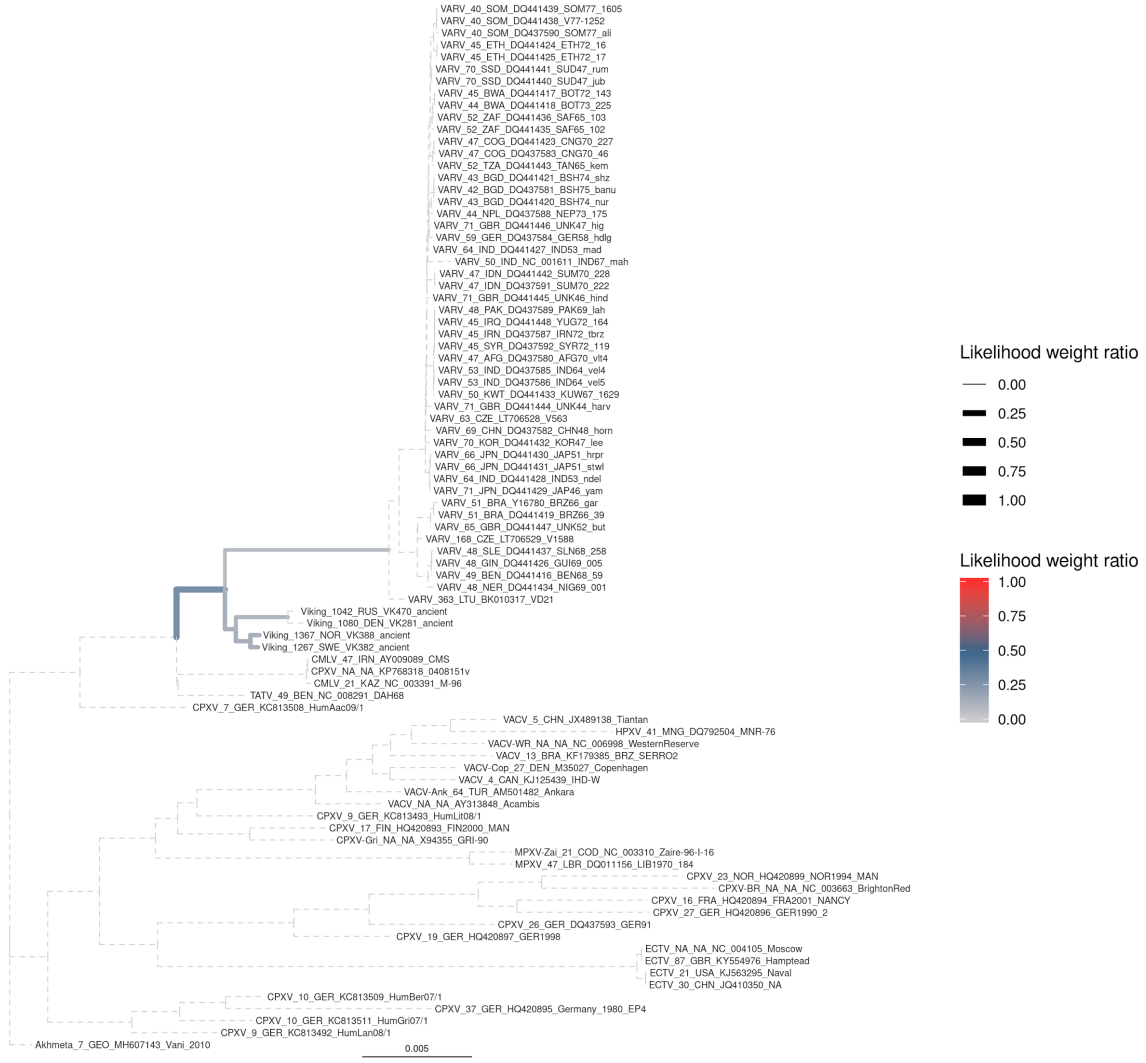

VARV_40_SOM_DQ441439_SOM77_1605
VARV_40_SOM_DQ441438_V77-1252
VARV_40_SOM_DQ437590_SOM77_ali
VARV_45_ETH_DQ441424_ETH72_16
VARV_45_ETH_DQ441425_ETH72_17
VARV_70_SSD_DQ441441_SUD47_rum
VARV_70_SSD_DQ441440_SUD47_jub
VARV_45_BWA_DQ441417_BOT72_143
VARV_44_BWA_DQ441418_BOT73_225
VARV_52_ZAF_DQ441436_SAF65_103
VARV_52_ZAF_DQ441435_SAF65_102
VARV_47_COG_DQ441423_CNG70_227
VARV_47_COG_DQ437583_CNG70_46
VARV_52_TZA_DQ441443_TAN65_kem
VARV_43_BGD_DQ441421_BSH74_shz
VARV_42_BGD_DQ437581_BSH75_banu
VARV_43_BGD_DQ441420_BSH74_nur
VARV_44_NPL_DQ437588_NEP73_175
VARV_71_GBR_DQ441446_UNK47_hig
VARV_59_GER_DQ437584_GER58_hdlg
VARV_64_IND_DQ441427_IND53_mad
VARV_50_IND_NC_001611_IND67_mah
VARV_47_IDN_DQ441442_SUM70_228
VARV_47_IDN_DQ437591_SUM70_222
VARV_71_GBR_DQ441445_UNK46_hind
VARV_48_PAK_DQ437589_PAK69_lah
VARV_45_IRQ_DQ441448_YUG72_164
VARV_45_IRN_DQ437587_IRN72_tbrz
VARV_45_SYR_DQ437592_SYR72_119
VARV_47_AFG_DQ437580_AFG70_vlt4
VARV_53_IND_DQ437585_IND64_vel4
VARV_53_IND_DQ437586_IND64_vel5
VARV_50_KWT_DQ441433_KUW67_1629
VARV_71_GBR_DQ441444_UNK44_harv
VARV_63_CZE_LT706528_V563
VARV_69_CHN_DQ437582_CHN48_horn
VARV_70_KOR_DQ441432_KOR47_lee
VARV_66_JPN_DQ441430_JAP51_hrpr
VARV_66_JPN_DQ441431_JAP51_stwl
VARV_64_IND_DQ441428_IND53_ndel
VARV_71_JPN_DQ441429_JAP46_yam
VARV_51_BRA_Y16780_BRZ66_gar
VARV_51_BRA_DQ441419_BRZ66_39
VARV_65_GBR_DQ441447_UNK52_but
VARV_168_CZE_LT706529_V1588
VARV_48_SLE_DQ441437_SLN68_258
VARV_48_GIN_DQ441426_GUI69_005
VARV_49_BEN_DQ441416_BEN68_59
VARV_48_NER_DQ441434_NIG69_001
VARV_363_LTU_BK010317_VD21
Viking_1042_RUS_VK470_ancient
Viking_1080_DEN_VK281_ancient
Viking_1367_NOR_VK388_ancient
Viking_1267_SWE_VK382_ancient
CMLV_47_IRN_AY009089_CMS
CPXV_NA_NA_KP768318_0408151v
CMLV_21_KAZ_NC_003391_M-96
TATV_49_BEN_NC_008291_DAH68
CPXV_7_GER_KC813508_HumAac09/1
VACV_5_CHN_JX489138_Tiantan
HPXV_41_MNG_DQ792504_MNR-76
VACV-WR_NA_NA_NC_006998_WesternReserve
VACV_13_BRA_KF179385_BRZ_SERRO2
VACV-Cop_27_DEN_M35027_Copenhagen
VACV_4_CAN_KJ125439_IHD-W
VACV-Ank_64_TUR_AM501482_Ankara
VACV_NA_NA_AY313848_Acambis
CPXV_9_GER_KC813493_HumLit08/1
CPXV_17_FIN_HQ420893_FIN2000_MAN
CPXV-Gri_NA_NA_X94355_GRI-90
MPXV-Zai_21_COD_NC_003310_Zaire-96-I-16
MPXV_47_LBR_DQ011156_LIB1970_184
CPXV_23_NOR_HQ420899_NOR1994_MAN
CPXV-BR_NA_NA_NC_003663_BrightonRed
CPXV_16_FRA_HQ420894_FRA2001_NANCY
CPXV_27_GER_HQ420896_GER1990_2
CPXV_26_GER_DQ437593_GER91
CPXV_19_GER_HQ420897_GER1998
ECTV_NA_NA_NC_004105_Moscow
ECTV_87_GBR_KY554976_Hamptead
ECTV_21_USA_KJ563295_Naval
ECTV_30_CHN_JQ410350_NA
CPXV_10_GER_KC813509_HumBer07/1
CPXV_37_GER_HQ420895_Germany_1980_EP4
CPXV_10_GER_KC813511_HumGri07/1
CPXV_9_GER_KC813492_HumLan08/1
Akhmeta_7_GEO_MH607143_Vani_2010
0.005
Likelihood weight ratio
0.00
0.25
0.50
0.75
1.00
Likelihood weight ratio
1.00
0.75
0.50
0.25
0.00

D

aVARV-VK515

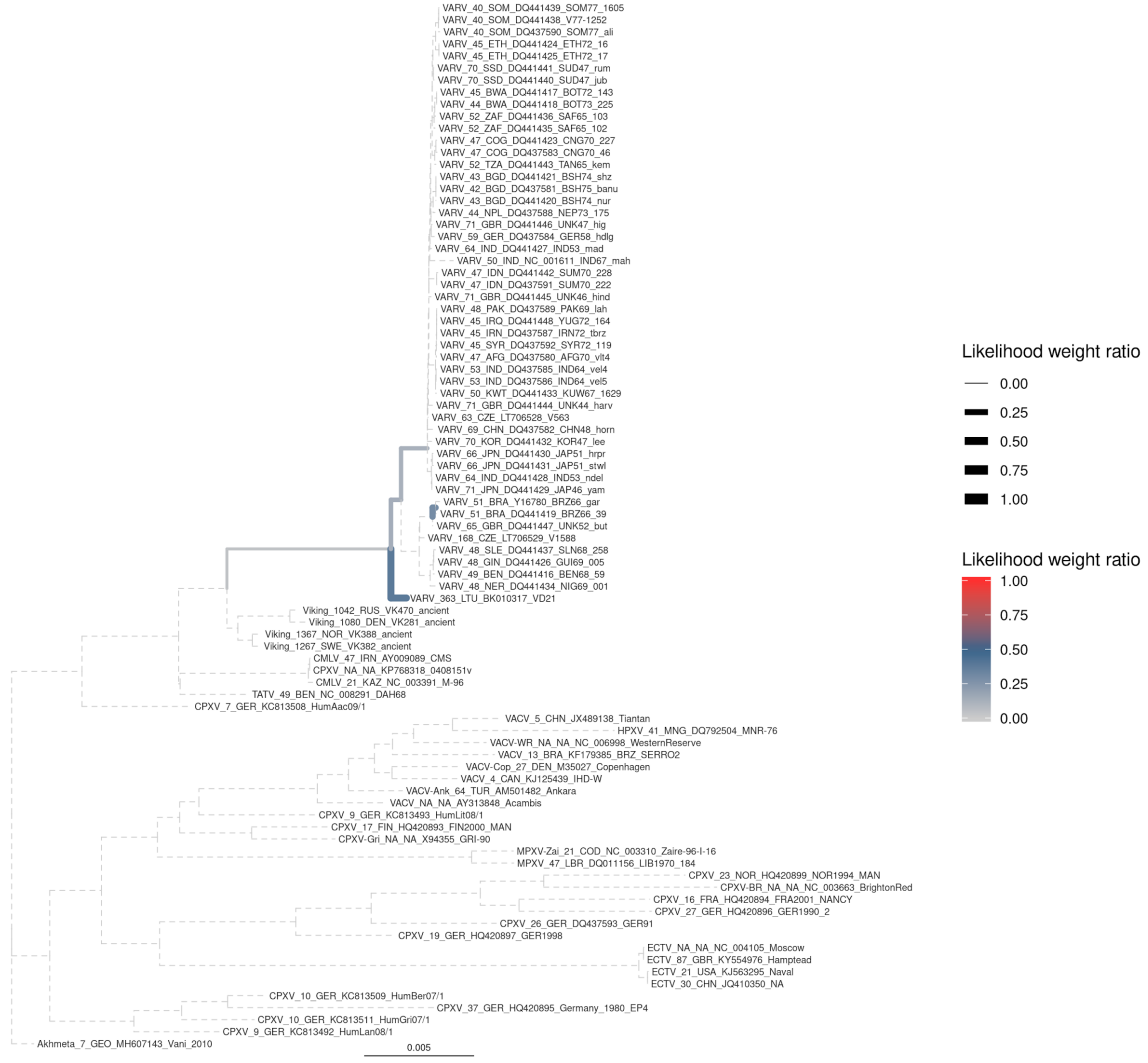
E

aVARV-VK443



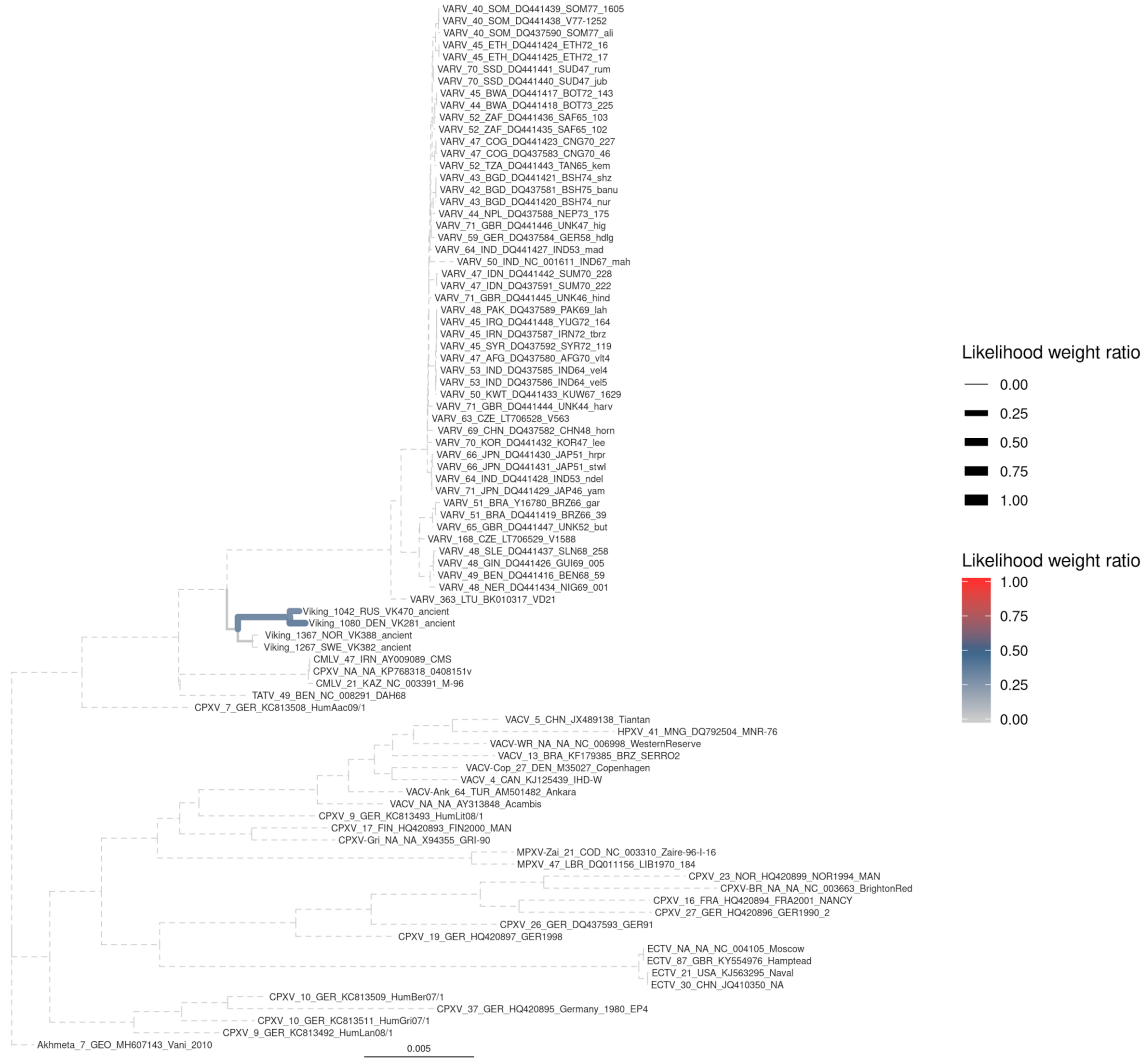
F

VARV-KHA1


VARV_40_SOM_DQ441439_SOM77_1605
VARV_40_SOM_DQ441438_V77-1252
VARV_40_SOM_DQ437590_SOM77_ali
VARV_45_ETH_DQ441424_ETH72_16
VARV_45_ETH_DQ441425_ETH72_17
VARV_70_SSD_DQ441441_SUD47_rum
VARV_70_SSD_DQ441440_SUD47_jub
VARV_45_BWA_DQ441417_BOT72_143
VARV_44_BWA_DQ441418_BOT73_225
VARV_52_ZAF_DQ441436_SAF65_103
VARV_52_ZAF_DQ441435_SAF65_102
VARV_47_COG_DQ441423_CNG70_227
VARV_47_COG_DQ437583_CNG70_46
VARV_52_TZA_DQ441443_TAN65_kem
VARV_43_BGD_DQ441421_BSH74_shz
VARV_42_BGD_DQ437581_BSH75_banu
VARV_43_BGD_DQ441420_BSH74_nur
VARV_44_NPL_DQ437588_NEP73_175
VARV_71_GBR_DQ441446_UNK47_hig
VARV_59_GER_DQ437584_GER58_hdlg
VARV_64_IND_DQ441427_IND53_mad
VARV_50_IND_NC_001611_IND67_mah
VARV_47_IDN_DQ441442_SUM70_228
VARV_47_IDN_DQ437591_SUM70_222
VARV_71_GBR_DQ441445_UNK46_hind
VARV_48_PAK_DQ437589_PAK69_lah
VARV_45_IRQ_DQ441448_YUG72_164
VARV_45_IRN_DQ437587_IRN72_tbrz
VARV_45_SYR_DQ437592_SYR72_119
VARV_47_AFG_DQ437580_AFG70_vlt4
VARV_53_IND_DQ437585_IND64_vel4
VARV_53_IND_DQ437586_IND64_vel5
VARV_50_KWT_DQ441433_KUW67_1629
VARV_71_GBR_DQ441444_UNK44_harv
VARV_63_CZE_LT706528_V563
VARV_69_CHN_DQ437582_CHN48_horn
VARV_70_KOR_DQ441432_KOR47_lee
VARV_66_JPN_DQ441430_JAP51_hrpr
VARV_66_JPN_DQ441431_JAP51_stwl
VARV_64_IND_DQ441428_IND53_ndel
VARV_71_JPN_DQ441429_JAP46_yam
VARV_51_BRA_Y16780_BRZ66_gar
VARV_51_BRA_DQ441419_BRZ66_39
VARV_65_GBR_DQ441447_UNK52_but
VARV_168_CZE_LT706529_V1588
VARV_48_SLE_DQ441437_SLN68_258
VARV_48_GIN_DQ441426_GUI69_005
VARV_49_BEN_DQ441416_BEN68_59
VARV_48_NER_DQ441434_NIG69_001
VARV_363_LTU_BK010317_VD21
Viking_1042_RUS_VK470_ancient
Viking_1080_DEN_VK281_ancient
Viking_1367_NOR_VK388_ancient
Viking_1267_SWE_VK382_ancient
CMLV_47_IRN_AY009089_CMS
CPXV_NA_NA_KP768318_0408151v
CMLV_21_KAZ_NC_003391_M-96
TATV_49_BEN_NC_008291_DAH68
CPXV_7_GER_KC813508_HumAac09/1
VACV_5_CHN_JX489138_Tiantan
HPXV_41_MNG_DQ792504_MNR-76
VACV-WR_NA_NA_NC_006998_WesternReserve
VACV_13_BRA_KF179385_BRZ_SERRO2
VACV-Cop_27_DEN_M35027_Copenhagen
VACV_4_CAN_KJ125439_IHD-W
VACV-Ank_64_TUR_AM501482_Ankara
VACV_NA_NA_AY313848_Acambis
CPXV_9_GER_KC813493_HumLit08/1
CPXV_17_FIN_HQ420893_FIN2000_MAN
CPXV-Gri_NA_NA_X94355_GRI-90
MPXV-Zai_21_COD_NC_003310_Zaire-96-I-16
MPXV_47_LBR_DQ011156_LIB1970_184
CPXV_23_NOR_HQ420899_NOR1994_MAN
CPXV-BR_NA_NA_NC_003663_BrightonRed
CPXV_16_FRA_HQ420894_FRA2001_NANCY
CPXV_27_GER_HQ420896_GER1990_2
CPXV_26_GER_DQ437593_GER91
CPXV_19_GER_HQ420897_GER1998
ECTV_NA_NA_NC_004105_Moscow
ECTV_87_GBR_KY554976_Hamptead
ECTV_21_USA_KJ563295_Naval
ECTV_30_CHN_JQ410350_NA
CPXV_10_GER_KC813509_HumBer07/1
CPXV_37_GER_HQ420895_Germany_1980_EP4
CPXV_10_GER_KC813511_HumGri07/1
CPXV_9_GER_KC813492_HumLan08/1
Akhmeta_7_GEO_MH607143_Vani_2010
0.005
Likelihood weight ratio
0.00
0.25
0.50
0.75
1.00
Likelihood weight ratio
1.00
0.75
0.50
0.25
0.00

G

aVARV-VK108



H

aVARV-VK138

I

aVARV-VK255

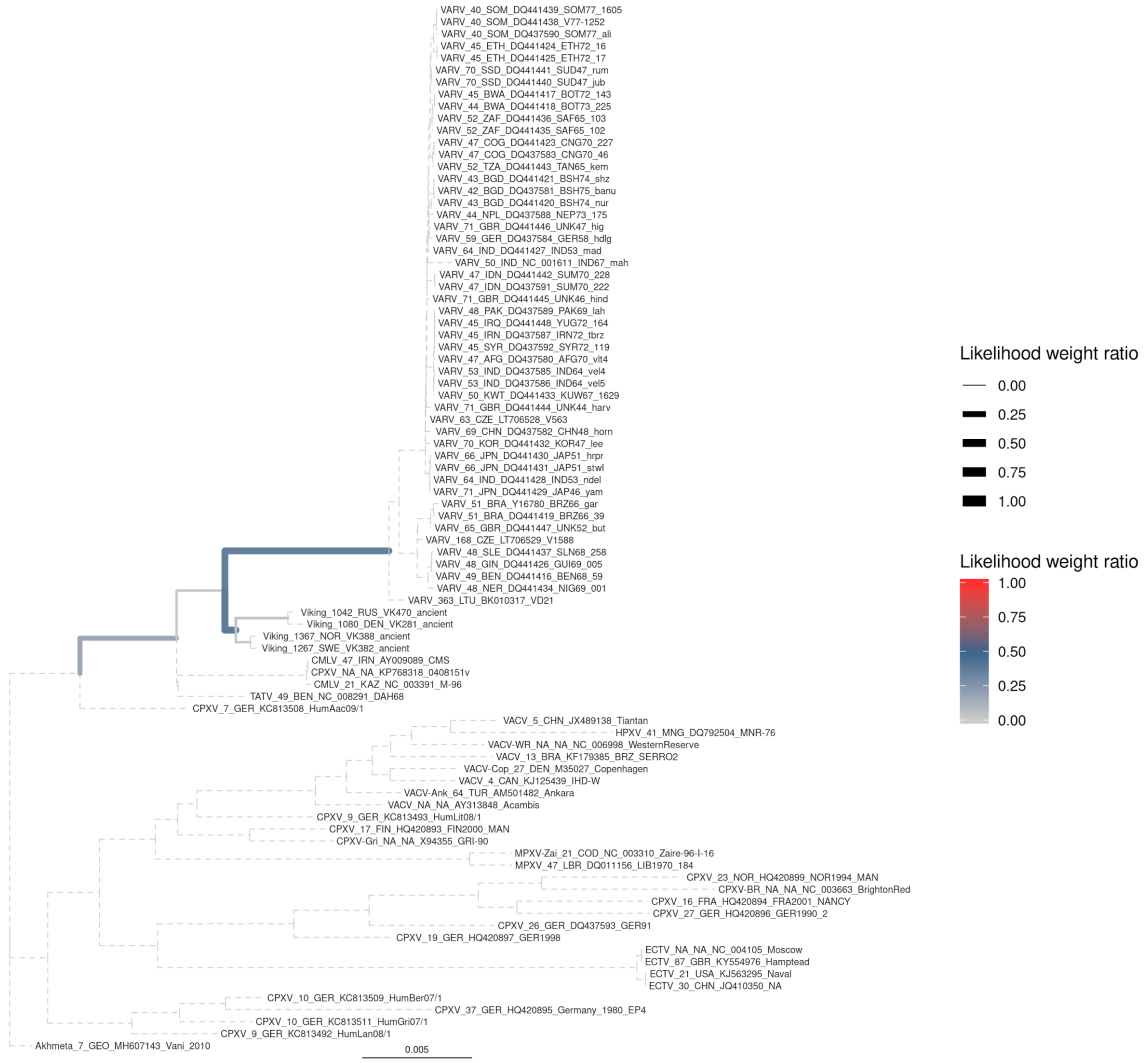


Fig. S7A-I. EPA placements of nine lower-coverage samples using EPA-ng. A) aVARV-VK168, B) aVARV-VK533, C) VARV-FIN1, D) aVARV-VK515, E) aVARV-VK443, F) VARV-KHA1, G) aVARV-VK108, H) aVARV-VK138, I) aVARV-VK255. Consensus sequences constructed from the reads from the lower-coverage samples were placed on a pre-computed ML tree (Fig. S6) using EPA-ng. Line colour and thickness indicate the likelihood weight ratio of the placement, showing the most-likely placement of the consensus on the pre-computed tree. All trees are rooted with Akhmeta_7_GEO_MH607143_Vani_2010.

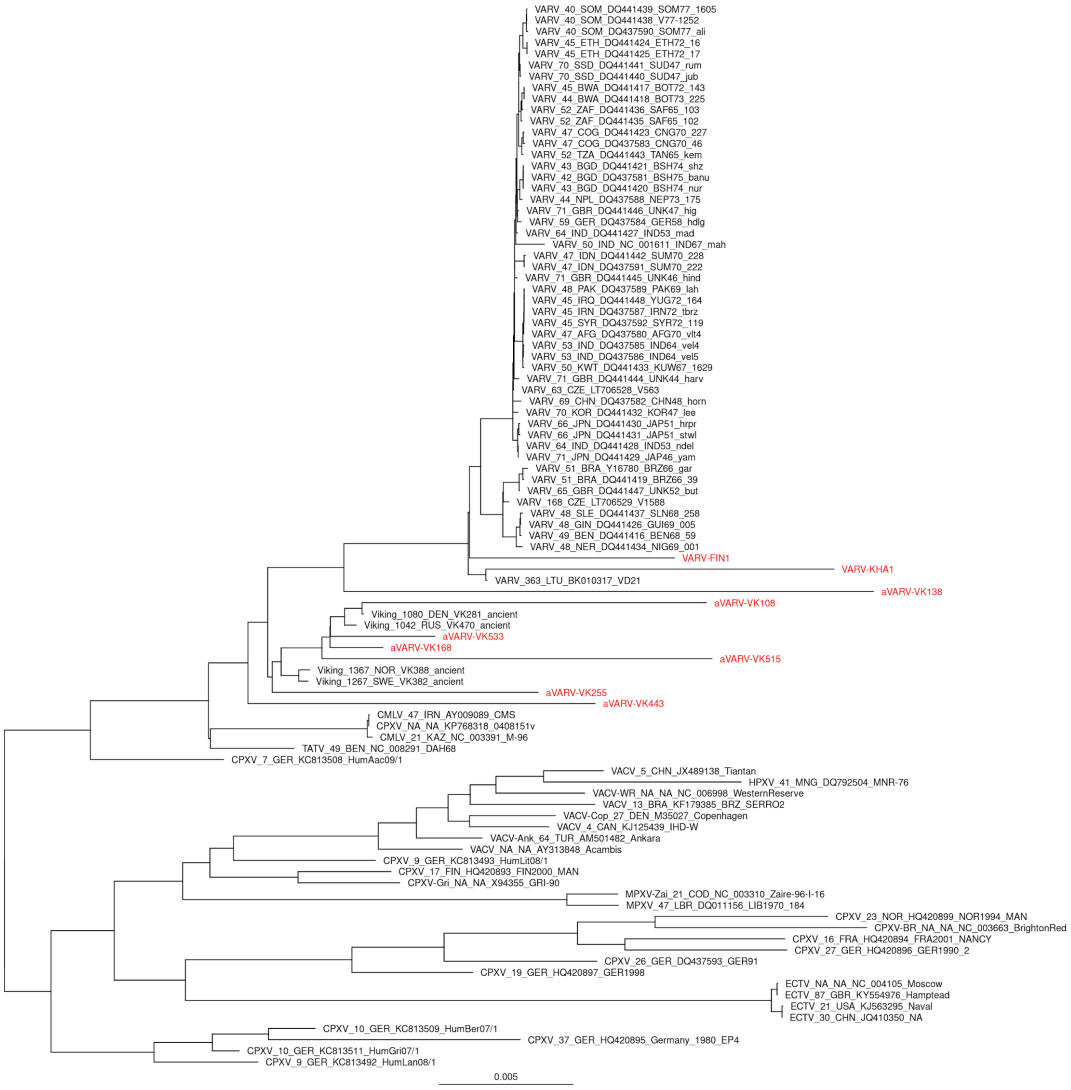
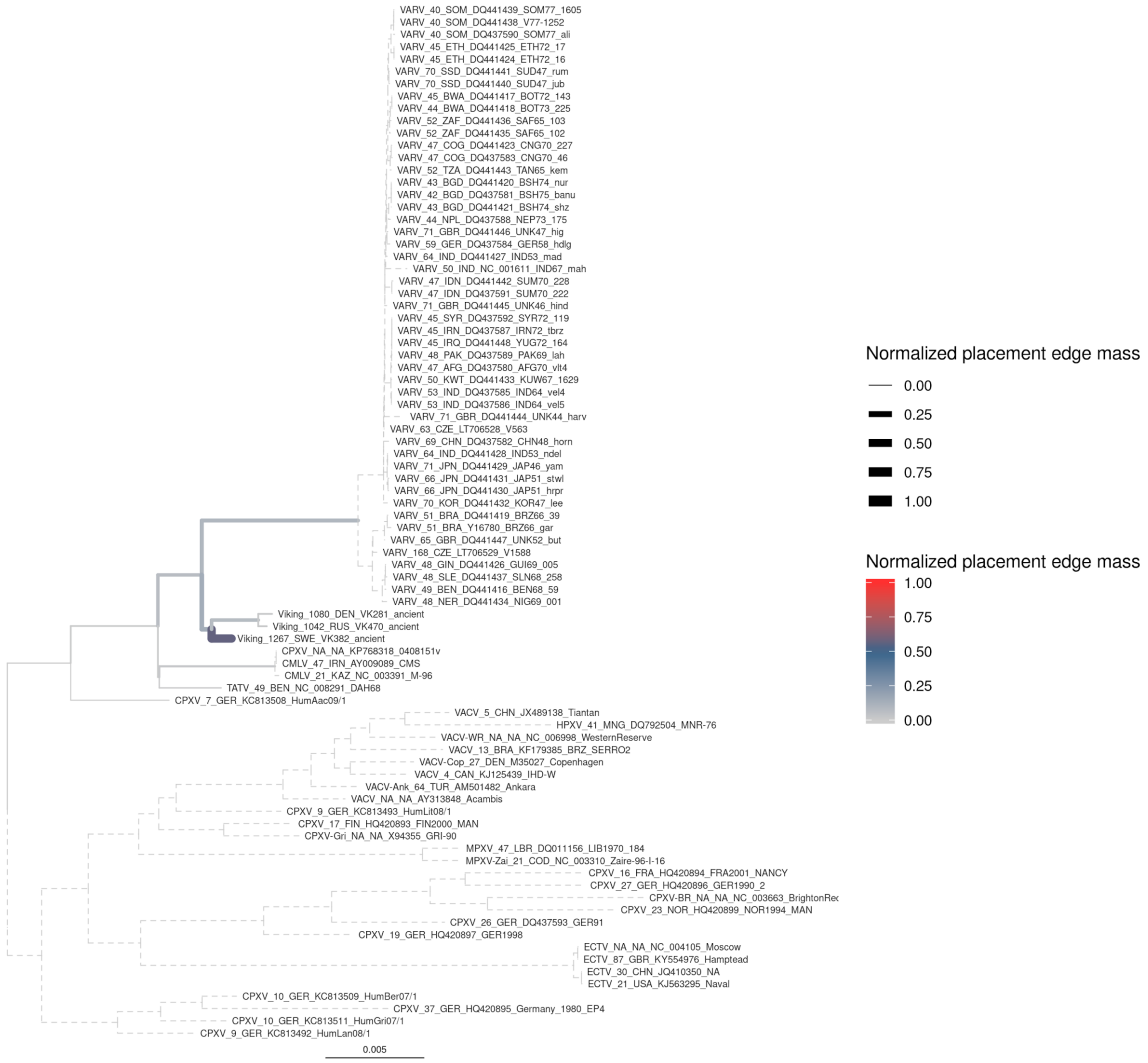


Fig. S8. Summary of the best EPA placements of all lower-coverage samples. Maximum likelihood tree showing the most-likely placements of the lower-coverage samples. Terminal branch lengths of lower-coverage samples are likely elongated because of increased inaccuracies in the consensus sequences due to low coverage depth.

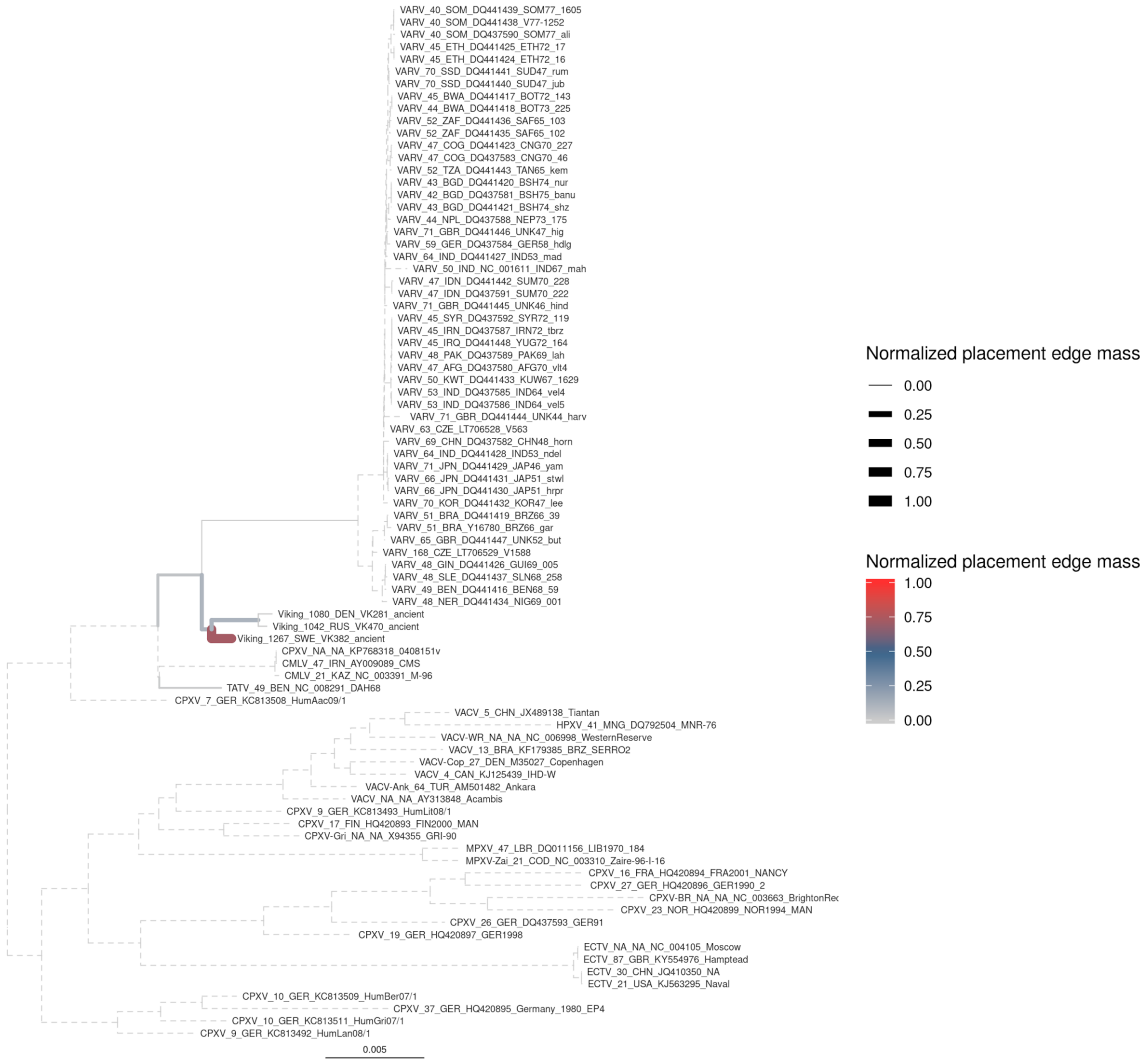
A

aVARV-VK388 / 30 reads subsample



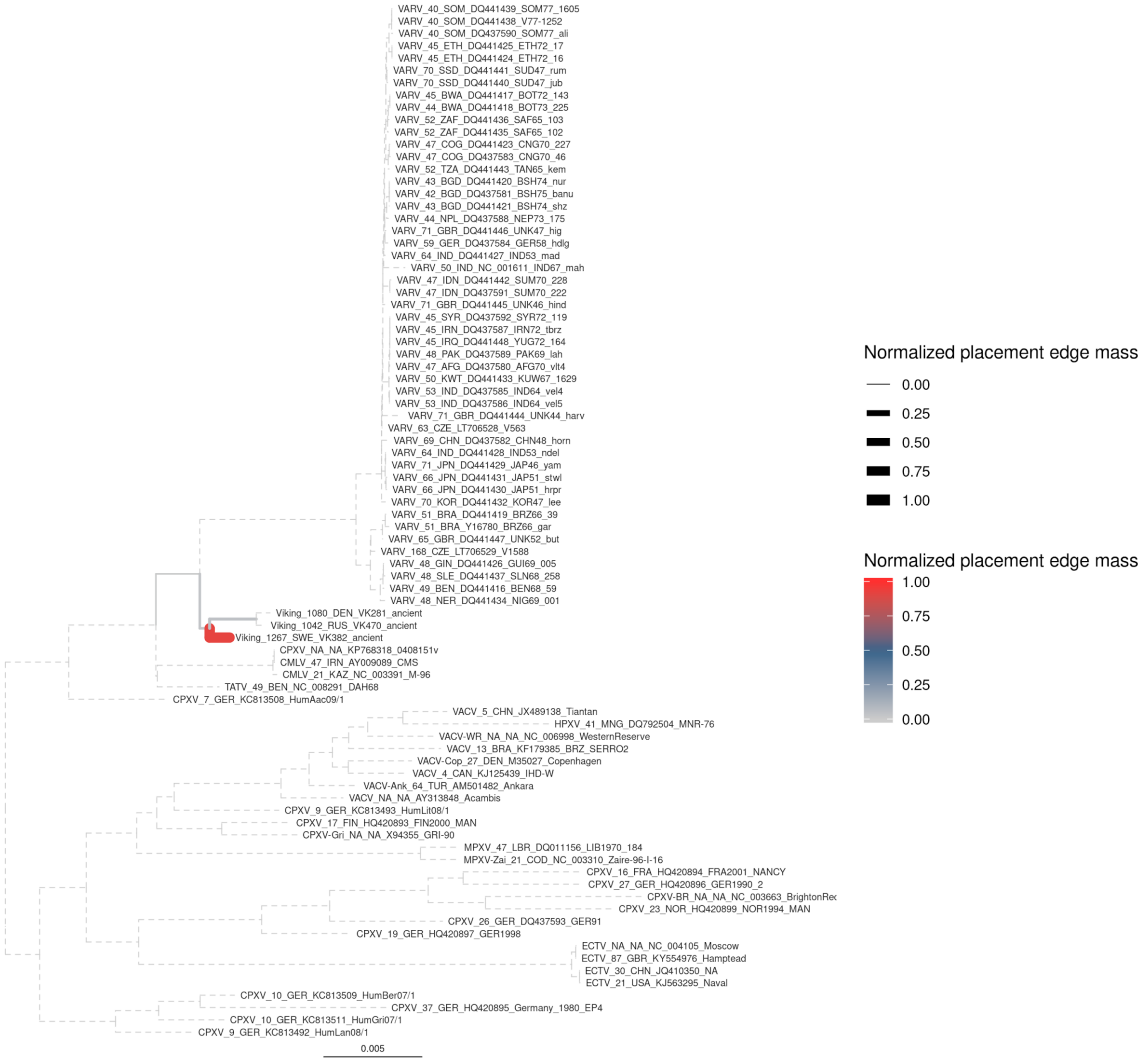
B

aVARV-VK388 / 50 reads subsample



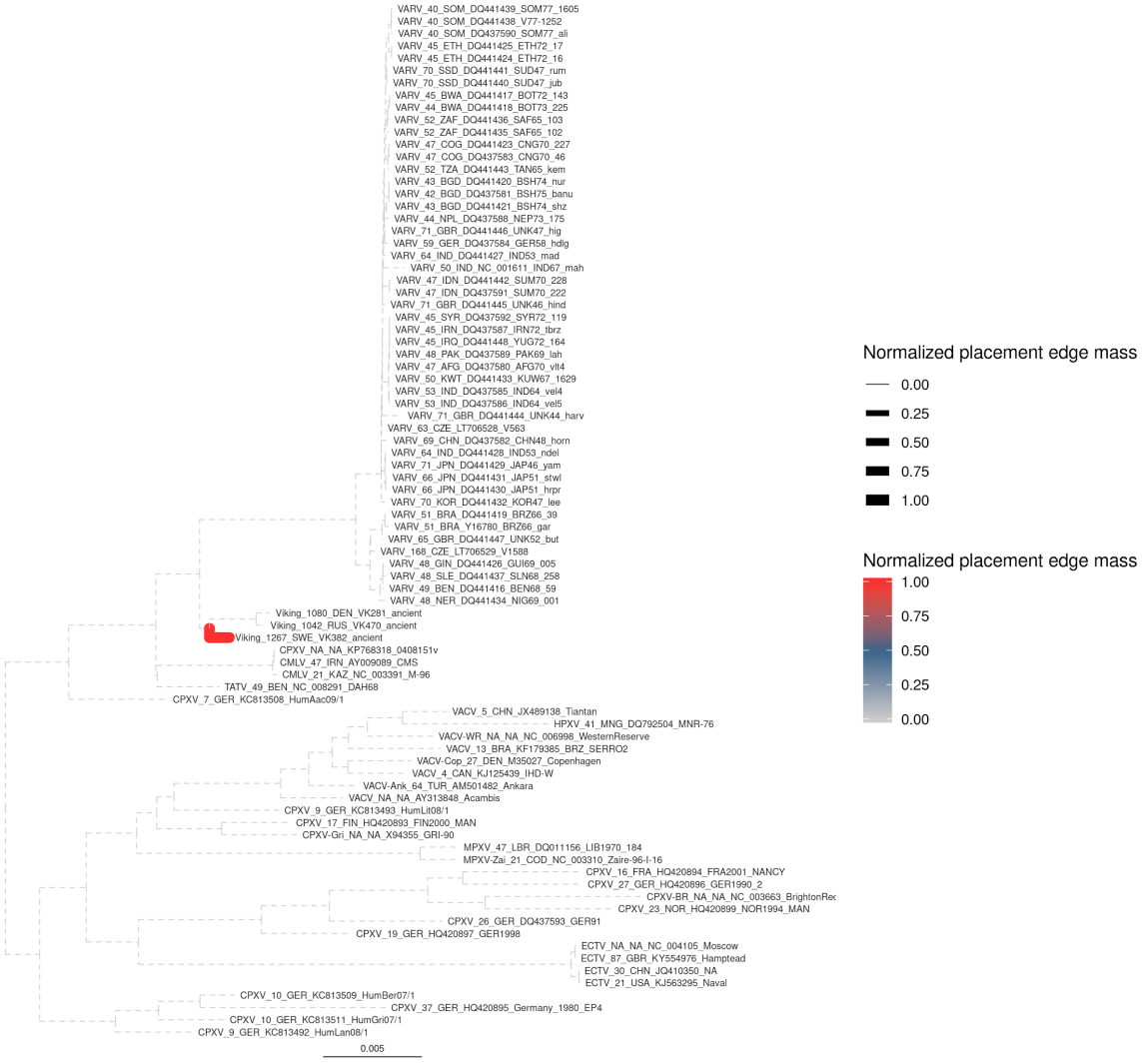
C

aVARV-VK388 / 100 reads subsample



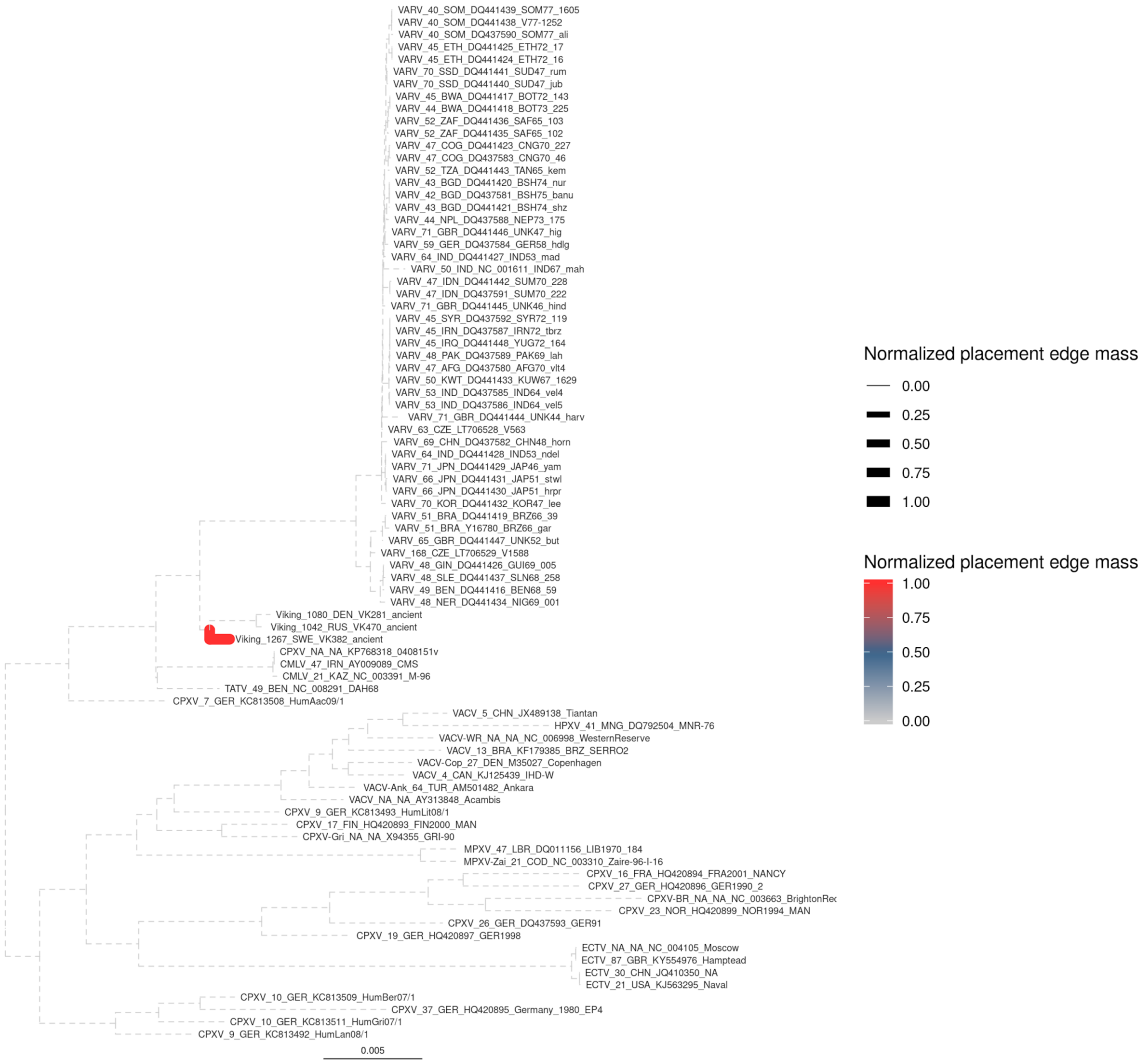
D

aVARV-VK388 / 500 reads subsample



E

aVARV-VK388 / 1000 reads subsample



F

aVARV-VK388 / 2000 reads subsample

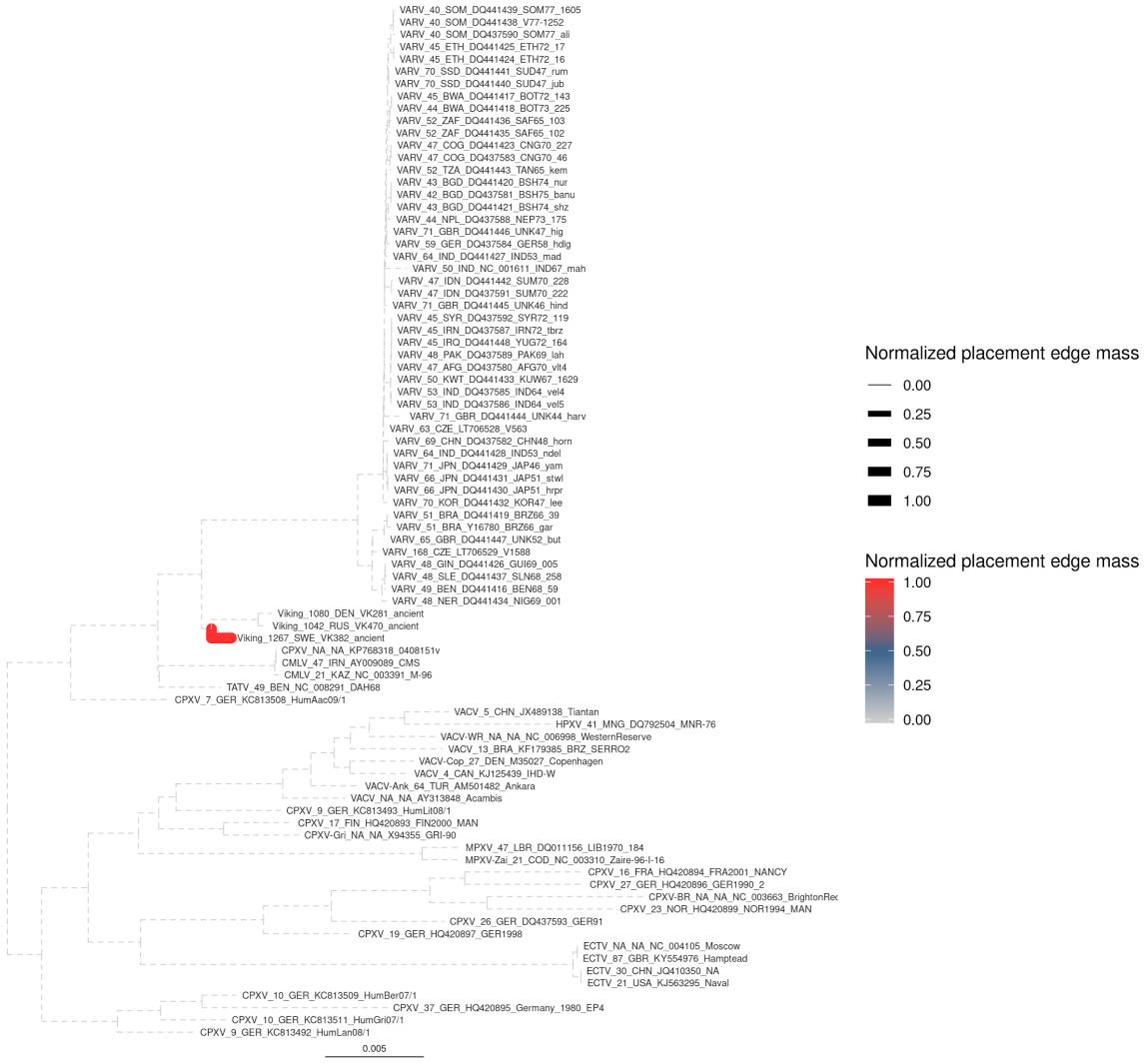
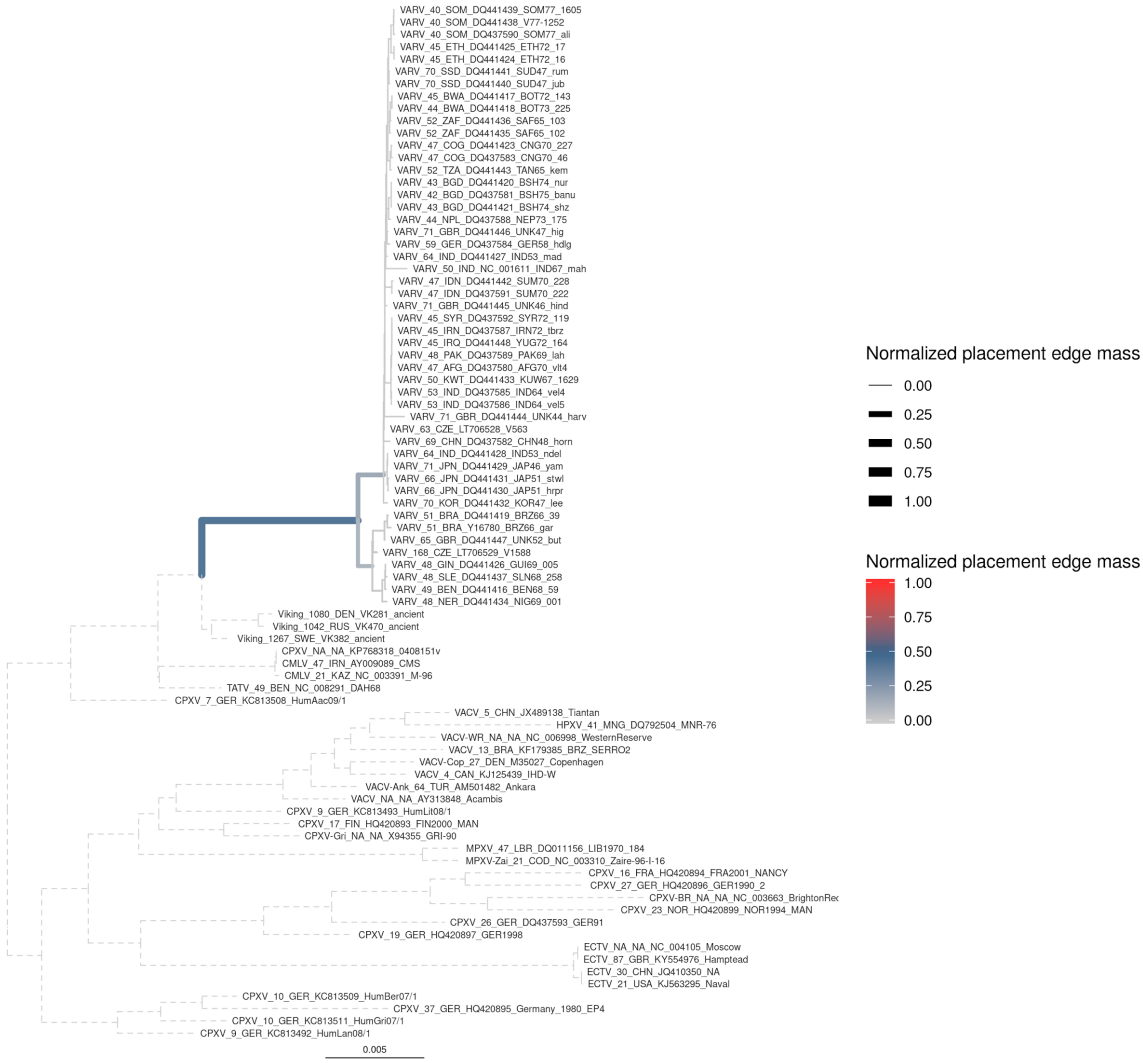


Fig. S9. EPA placements for random subsamples of reads from VK388. Consensus sequences generated from random subsample sets of reads from VK388 were placed onto a reference tree made without aVARV-VK388 and VARV-VD21 using EPA-ng. **A) 30 reads, B) 50 reads, C) 100 reads, D) 500 reads, E) 1000 reads, F) 2000 reads.** Coloured lines and line thickness indicate the normalize placement mass from 100 replicate consensus sequences for which the most-likely inferred placement was on the respective branch. All subsample sets showed the highest normalized placement mass on the expected branch (leading to aVARV-VK382), with sets with a higher number of reads (500 or more) placed with certainty (normalized placement edge mass = 1). Sets with the lowest number of reads (30 or 50) showed some uncertainty, however all placements were within the immediate neighbourhood of the aVARV clade, supporting the EPA-ng placement within a major clade (mVARV or aVARV) when run on samples with as few as 30 reads.

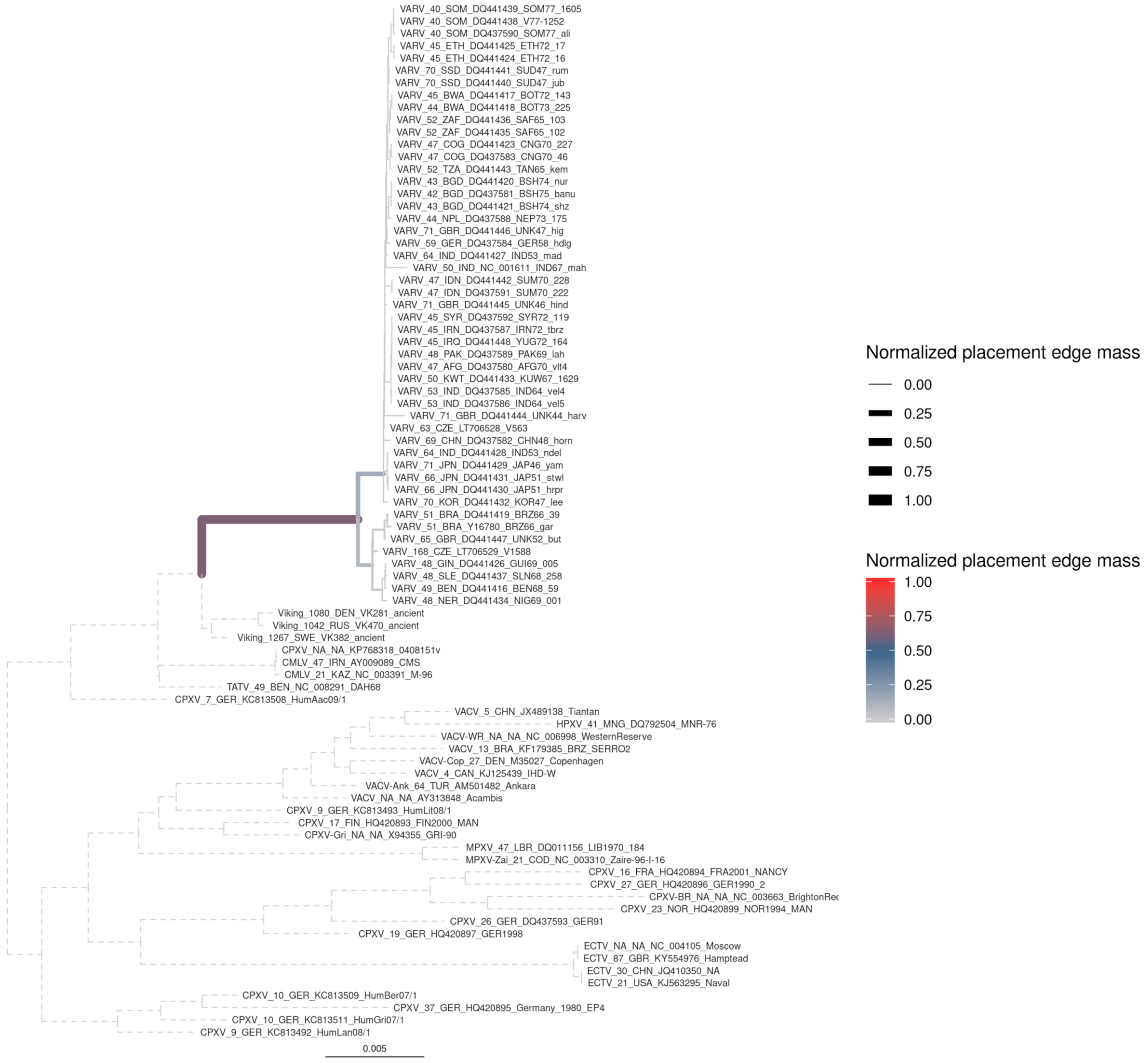
A

VD21 / 30 reads subsample



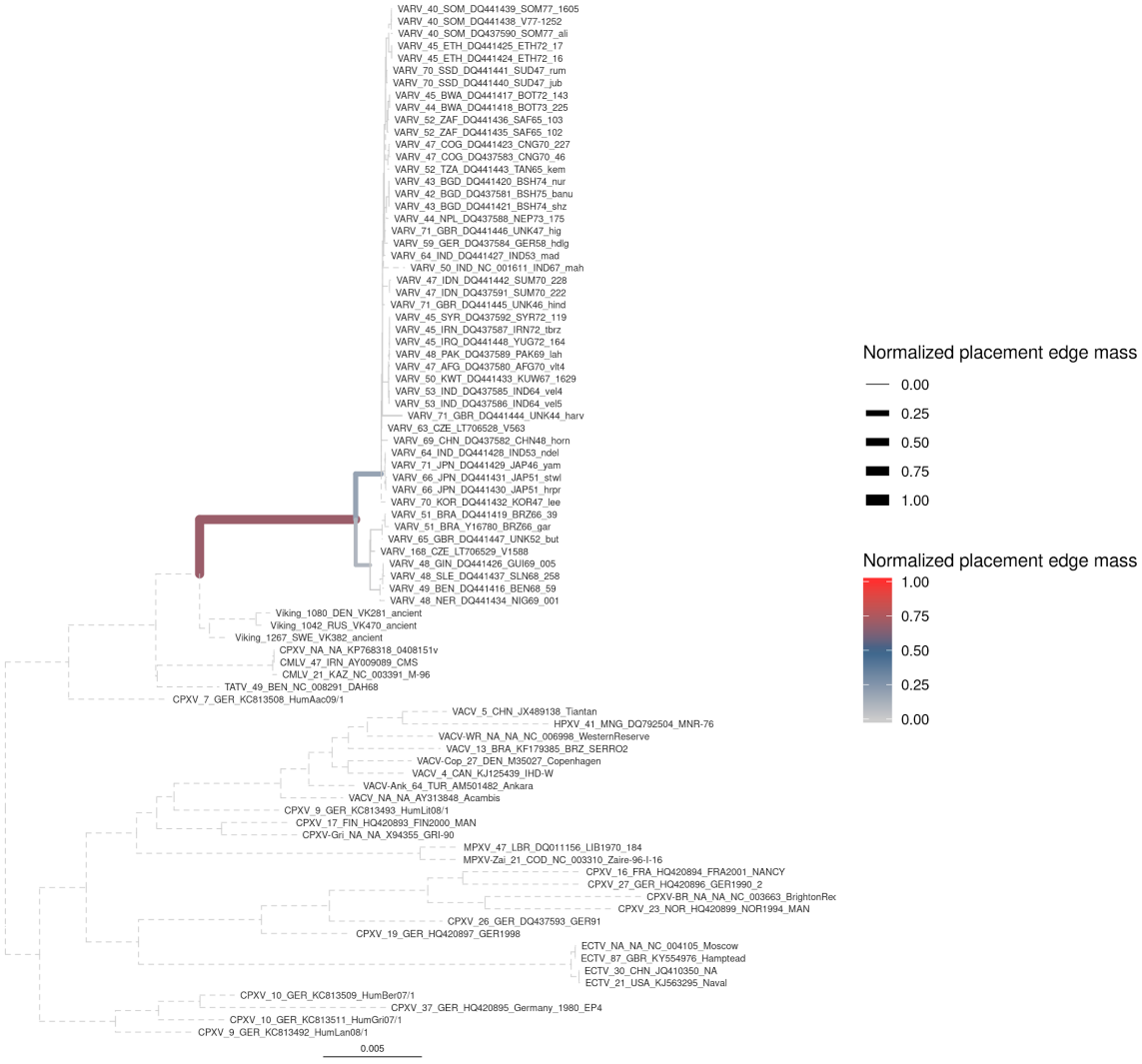
B

VD21 / 50 reads subsample



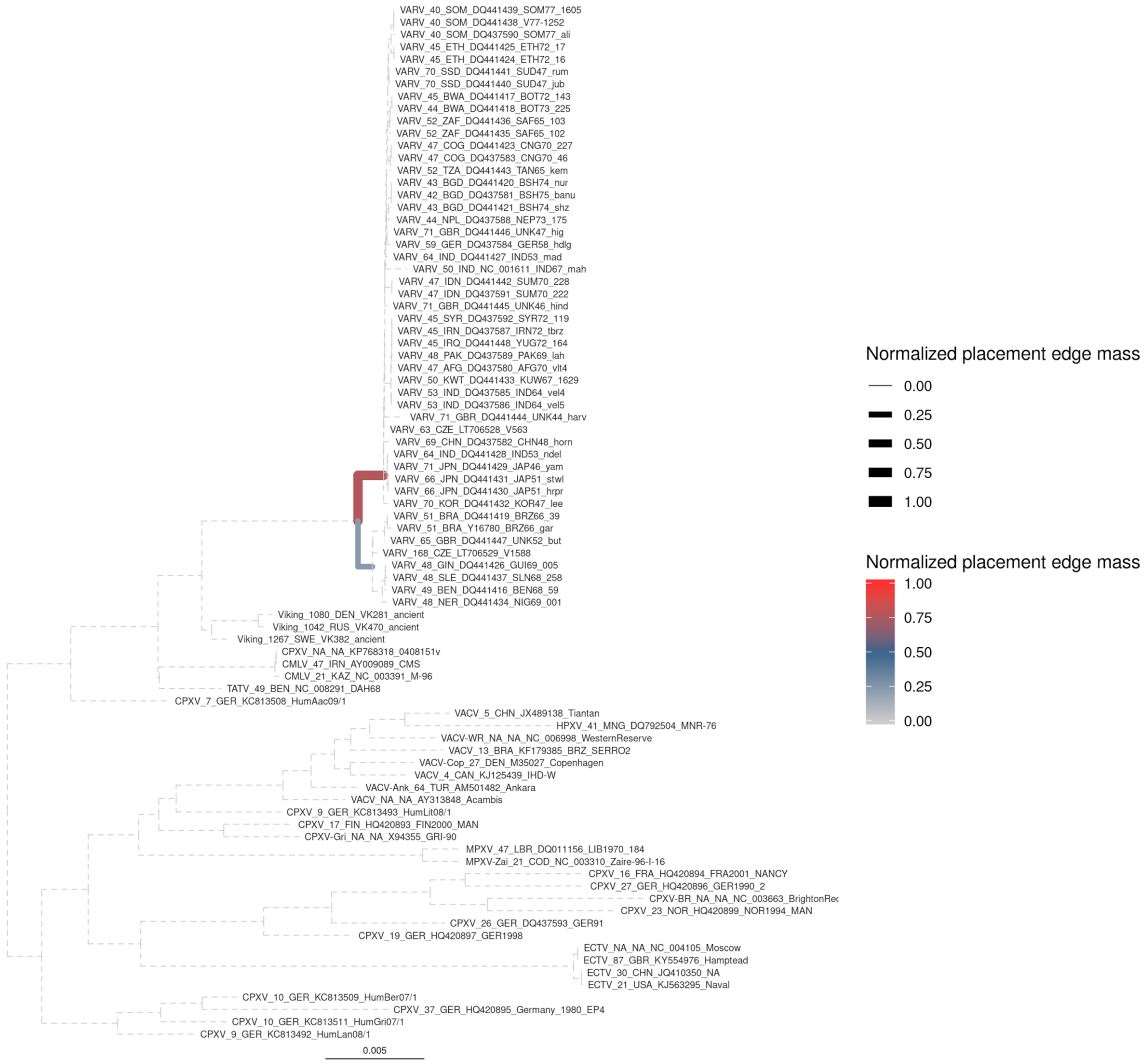
C

VD21 / 100 reads subsample



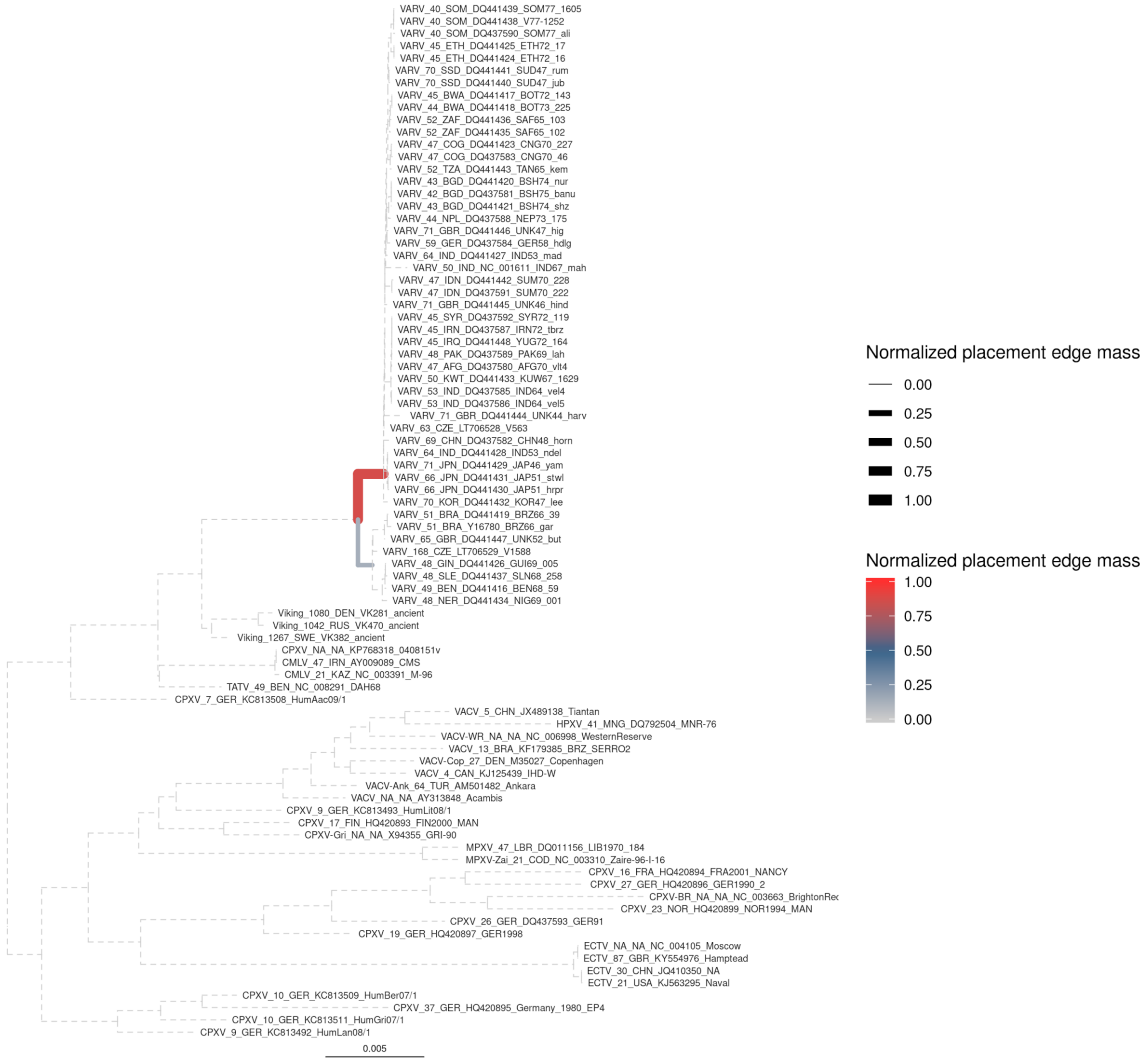
D

VD21 / 500 reads subsample



E

VD21 / 1000 reads subsample



F

VD21 / 2000 reads subsample

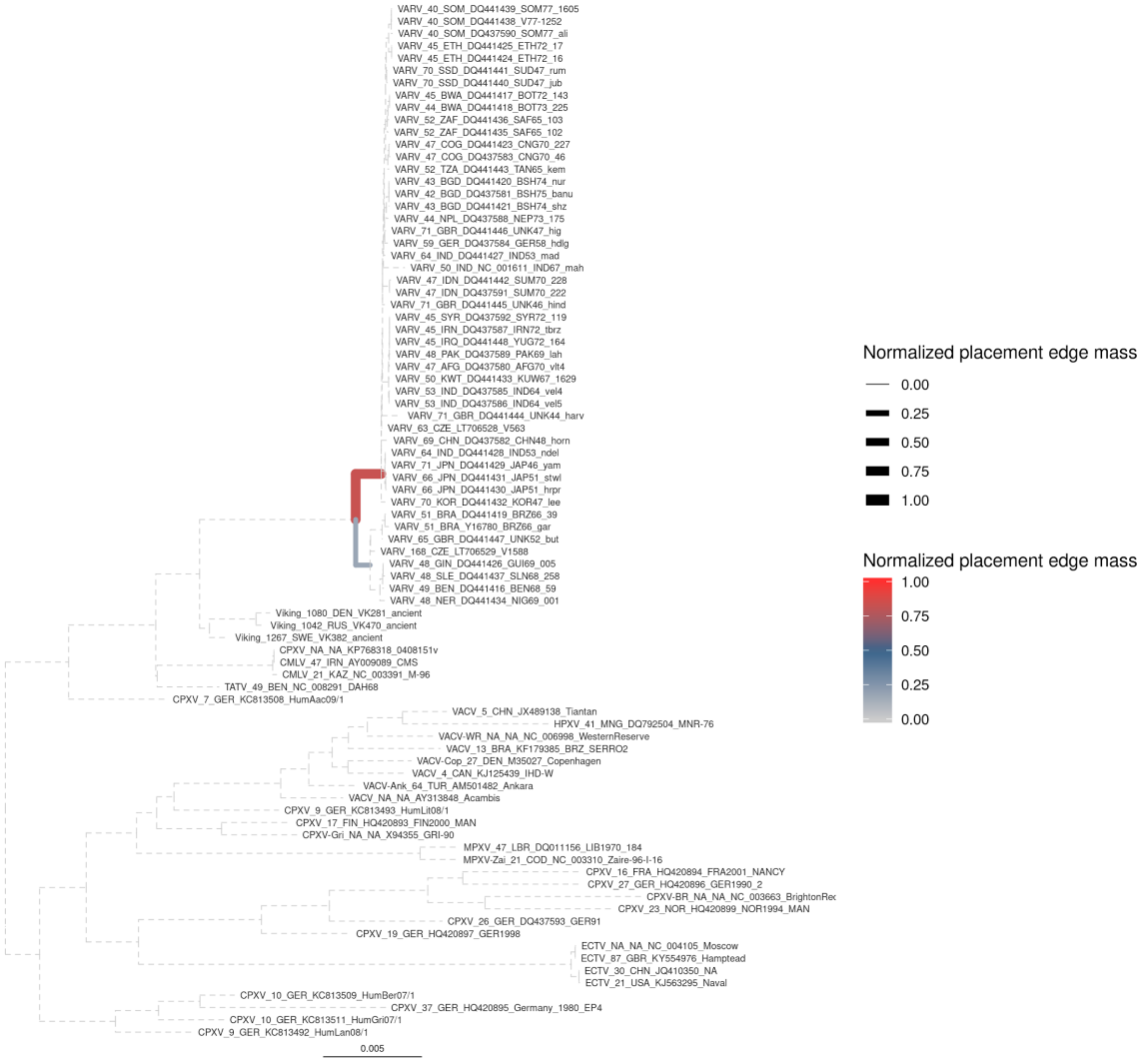


Fig. S10. EPA placements for random subsamples of reads from VD21. Consensus sequences generated from random subsample sets of reads from VD21 were placed onto a reference tree made without aVARV-VK388 and VARV-VD21 using EPA-ng. **A) 30 reads, B) 50 reads, C) 100 reads, D) 500 reads, E) 1000 reads, F) 2000 reads.** Coloured lines and line thickness indicate the normalize placement mass from 100 replicate consensus sequences for which the most-likely inferred placement was on the respective branch. As expected, all subsample sets showed the highest normalized placement mass on branches leading to or within the mVARV clade.

A

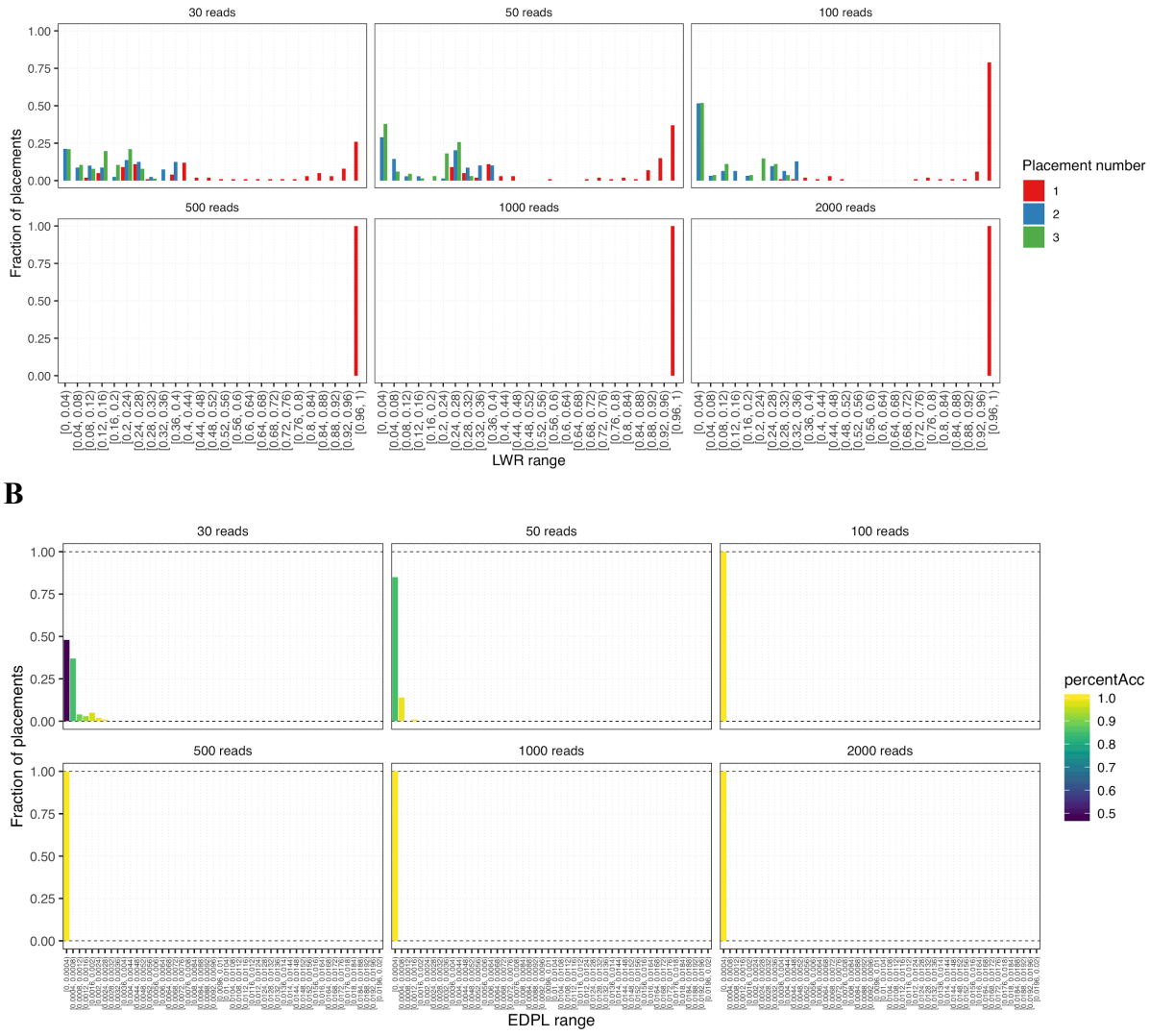
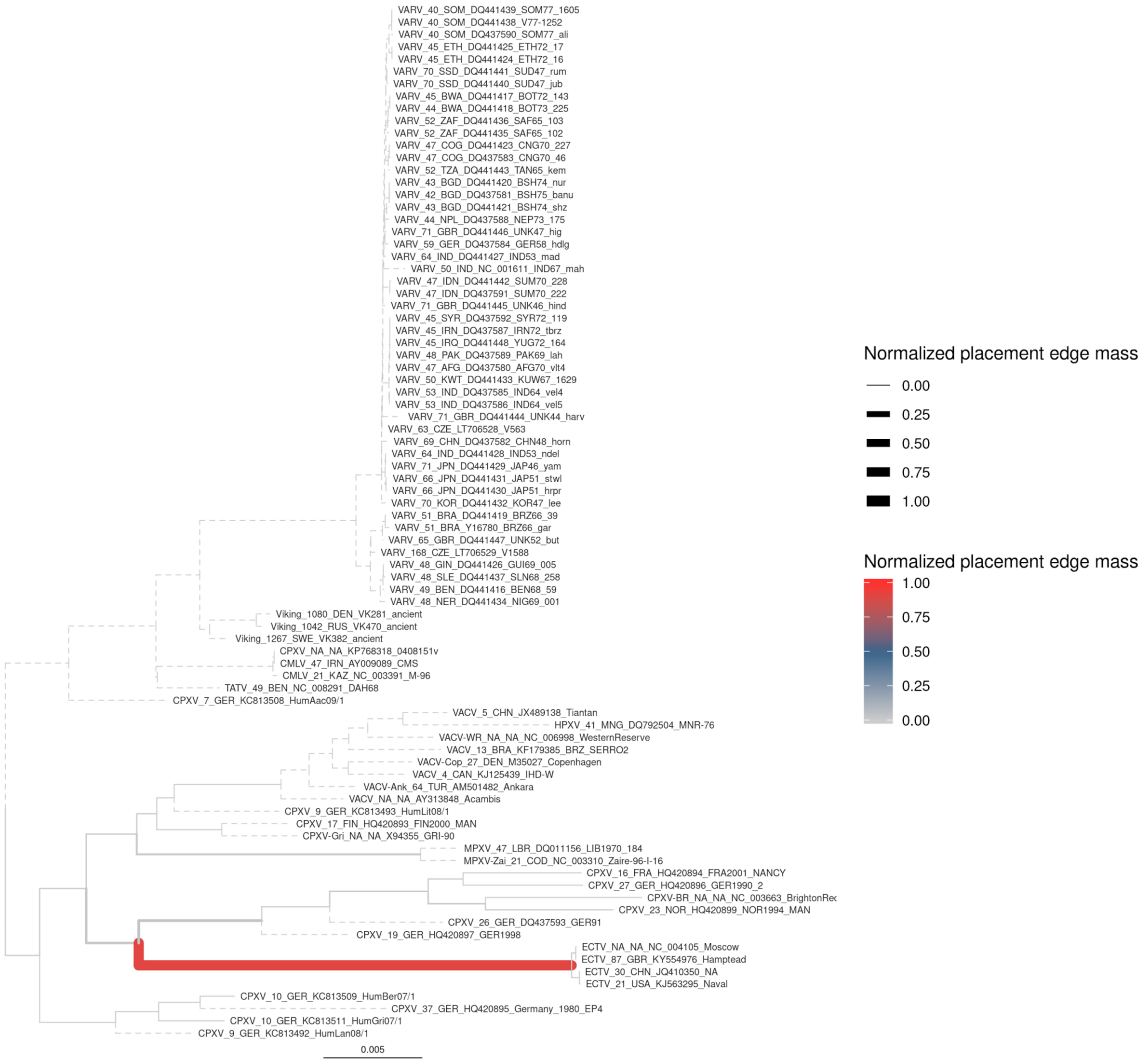


Fig. S11. EPA placement reliability. **A) Distributions of LWRs for the first, second, and third most-likely placement for each subsample set.** Sets with higher numbers of reads (≥ 500) show high confidence placements (LWR > 0.95) over all replicates, whereas lower read numbers increase uncertainty in the placements. **B) Distributions of EDPL scores.** Bars show the fraction of EDPL scores within the respective bins, with cumulative percentage represented with colors. For subsets with 100 reads or more, placements are highly localized, with EDPL values less than 0.0004 for all replicates (for comparison, observed branch length of aVARV-VK382 in the reference tree is 0.001631). Variation increases for lower read numbers, however the vast majority of replicates are still consistently placed within a localized neighborhood (EDPL < 0.0008) for sets with 30 or 50 reads.

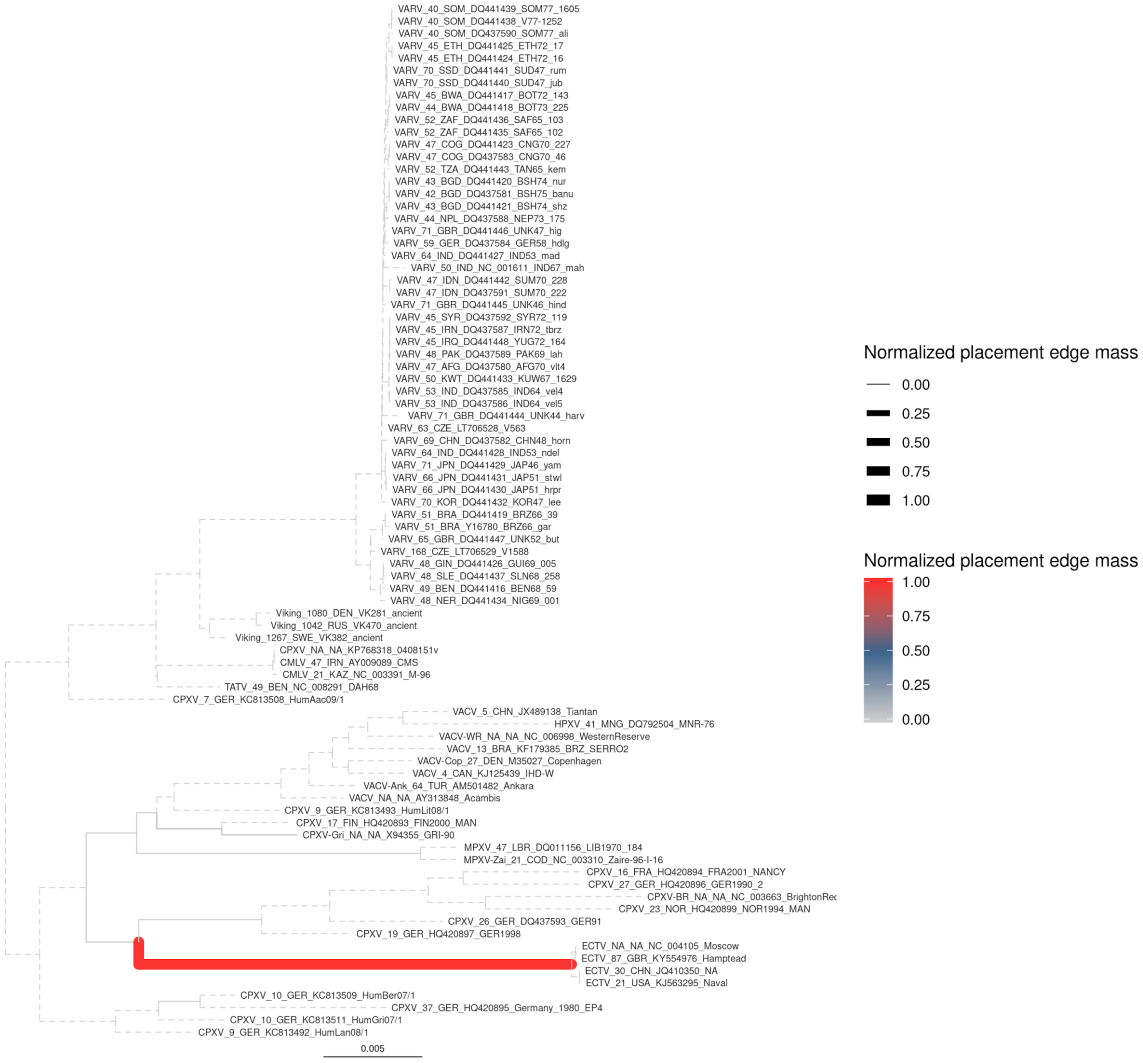
A

abatino_sim_1M / 30 reads subsample



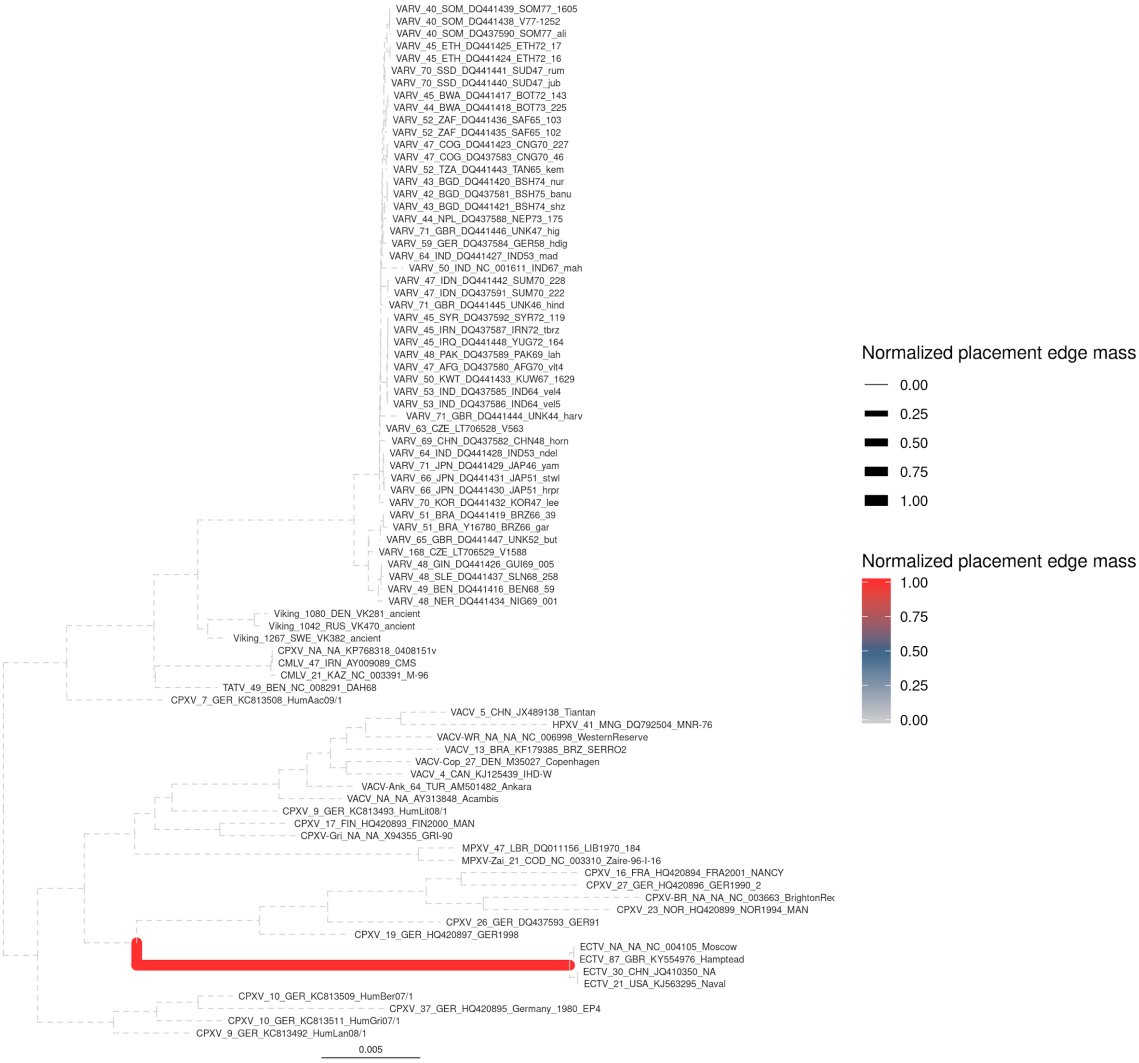
B

abatino_sim_1M / 50 reads subsample



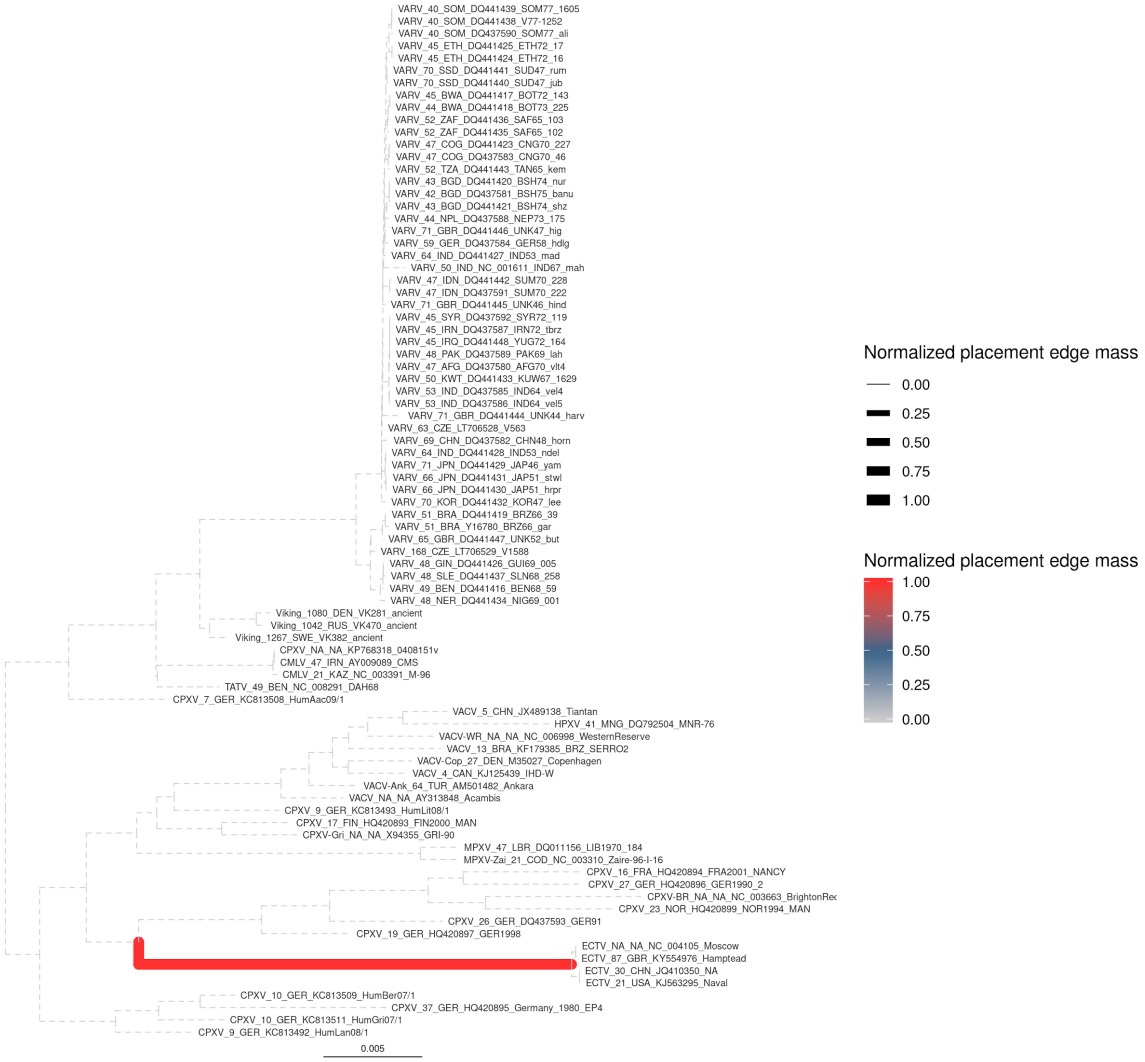
C

abatino_sim_1M / 100 reads subsample



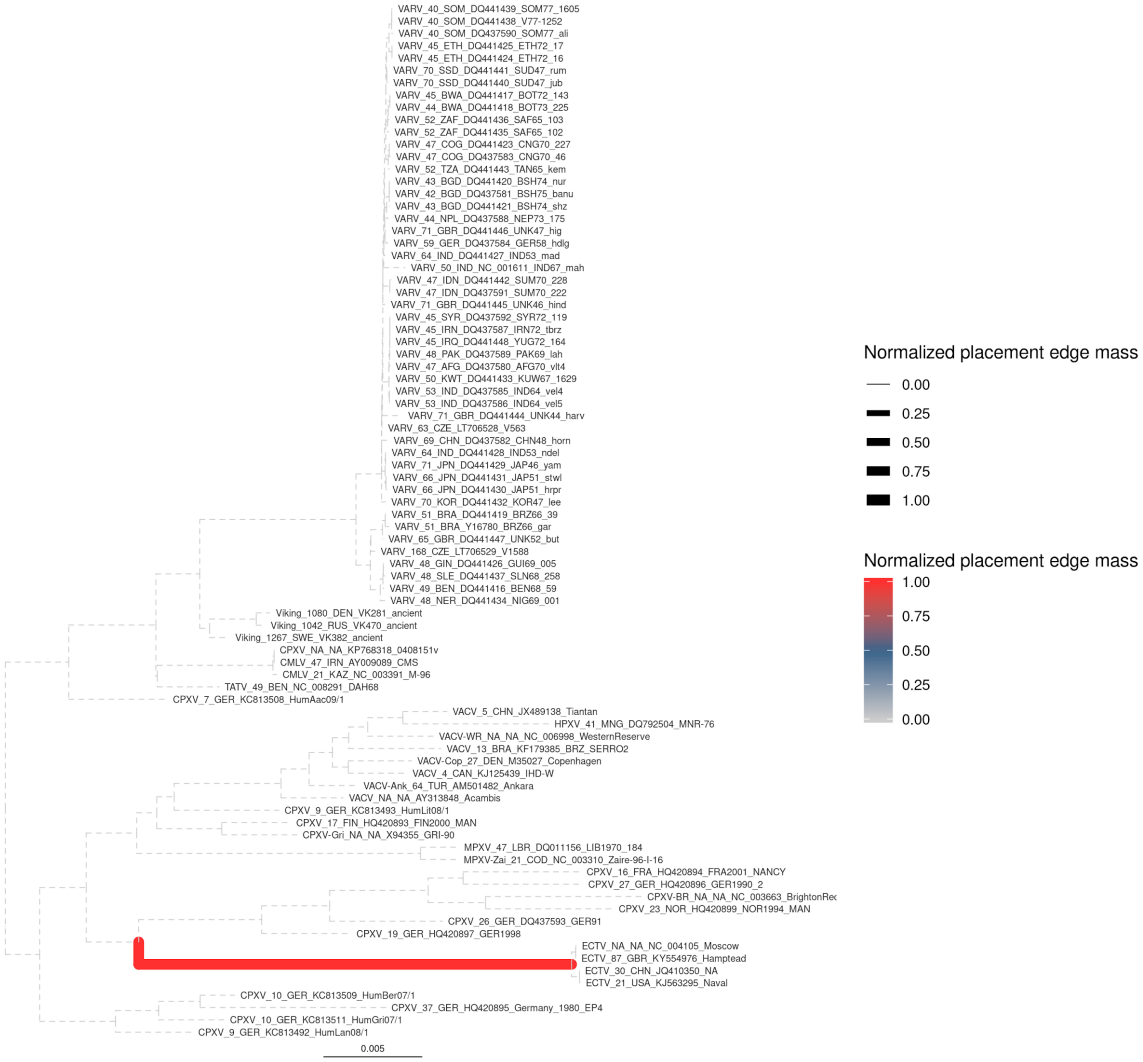
D

abatino_sim_1M / 500 reads subsample



E

abatino_sim_1M / 1000 reads subsample



F

abatino_sim_1M / 2000 reads subsample

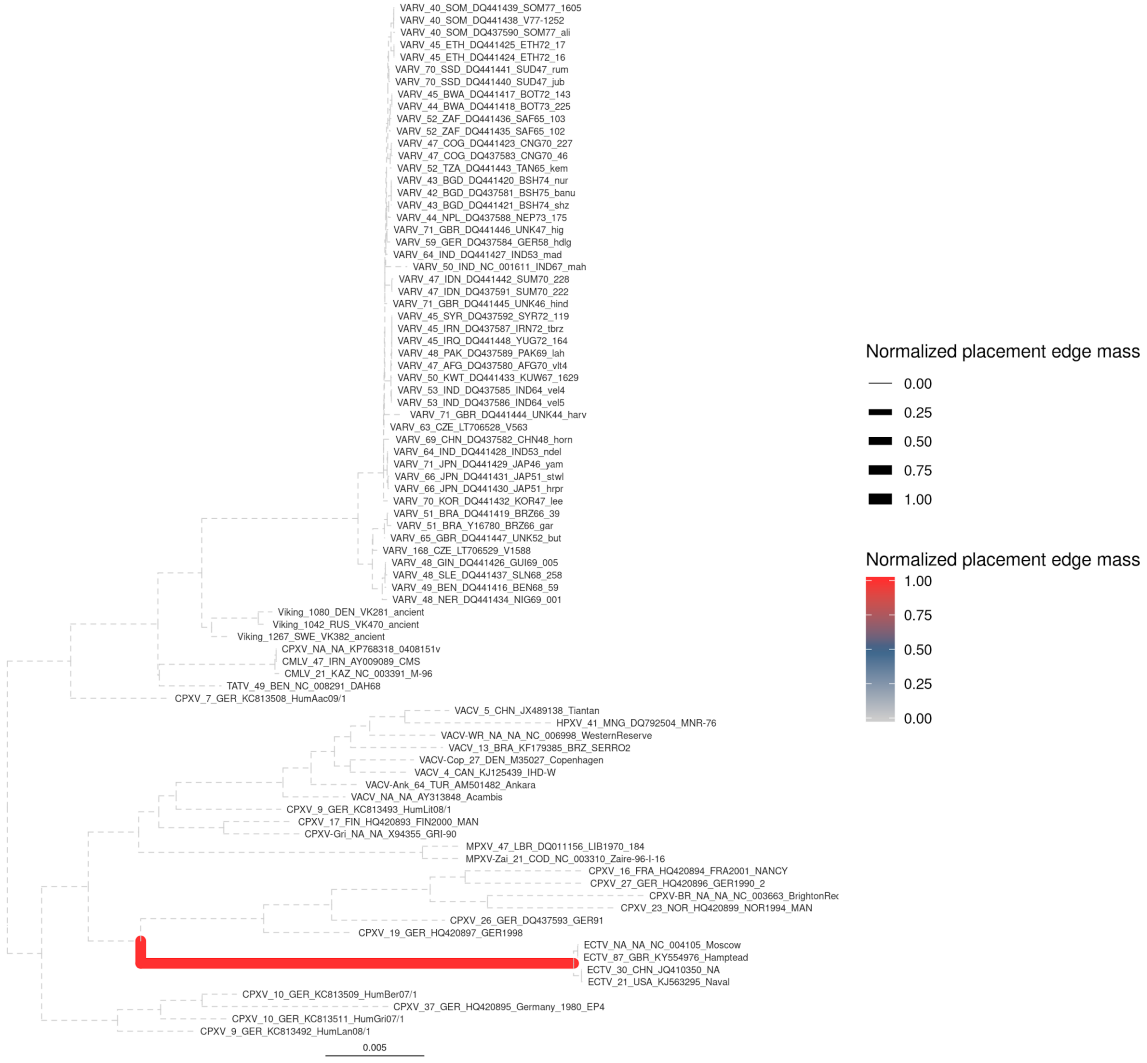


Fig. S12. EPA placements for random subsamples of simulated reads from orthopoxvirus Abatino. Consensus sequences generated from random subsample sets of reads simulated from the orthopoxvirus Abatino reference genome (GenBank accession no.: MH816996.1) were placed onto a reference tree using EPA-ng. **A) 30 reads, B) 50 reads, C) 100 reads, D) 500 reads, E) 1000 reads, F) 2000 reads.** Coloured lines and line thickness indicate the normalized placement mass fraction of from 100 replicate consensus sequences for which the most-likely inferred placement was on the respective branch.

Fig. S13. Maximum likelihood tree with higher-coverage samples used for the linear regression analysis (Fig. S14). Sequences were aligned using MAFFT. The tree was inferred using RAxML-NG with a K81uf+G+I substitution model, and a maximum likelihood estimate of tree topology, branch lengths, substitution rates, and nucleotide frequencies, with support estimated using 1000 bootstrap replicates. Branch lengths correspond to substitutions per site.

The aVARV sequences are shown in red, the ancient sequence from Duggan *et al.*, 2016 in blue (*16*). P-I and P-II refer to VARV primary clade I and primary clade II, respectively (*173*). 'alastrim' refers to alastrim minor, a phenotypically mild smallpox of America and West Africa (*173*). Taxon name fields indicate: virus/historical period, sample age relative to 2017, country abbreviation of sequence origin and region of sequence origin, as determined by the Standard country or area codes for statistical use, GenBank accession number/sample identifier, strain name. The x-axis denotes substitutions per site. The tree is rooted based on the position of the CMLV and TATV clade in Fig. S6.

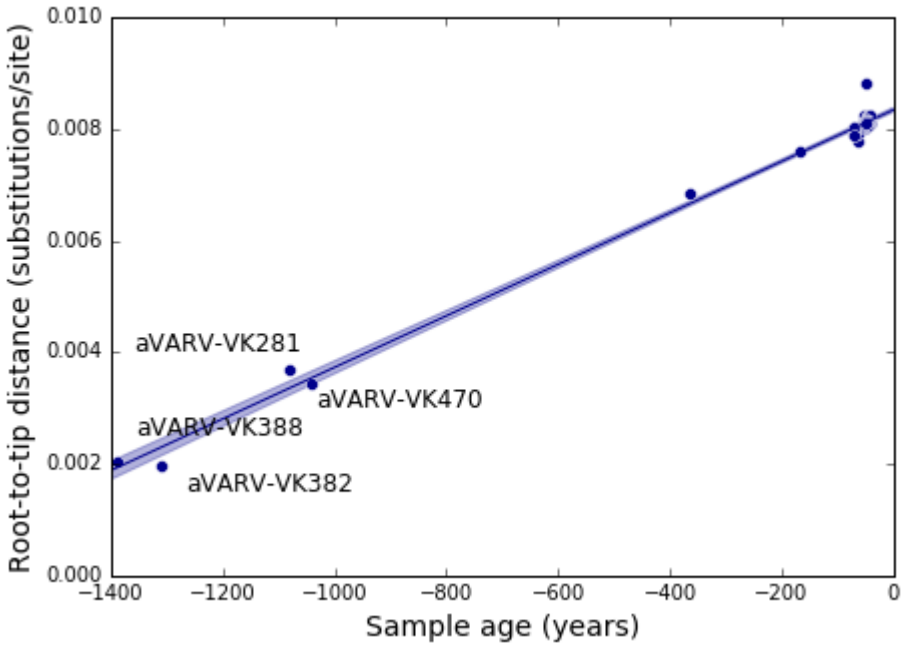


Fig. S14. Linear regression of root-to-tip distances against sampling dates. Root-to-tip distances were extracted from the ML tree shown in Fig. S13 using TempEst, using the 'heuristic residual mean squared' function and the 'best-fitting-root' option. The regression analysis was performed in SciPy. Shaded areas represent 95% confidence. Mean sample ages were used. The slope is 4.61×10^{-6} and the correlation coefficient is 0.99 (R^2=0.99).

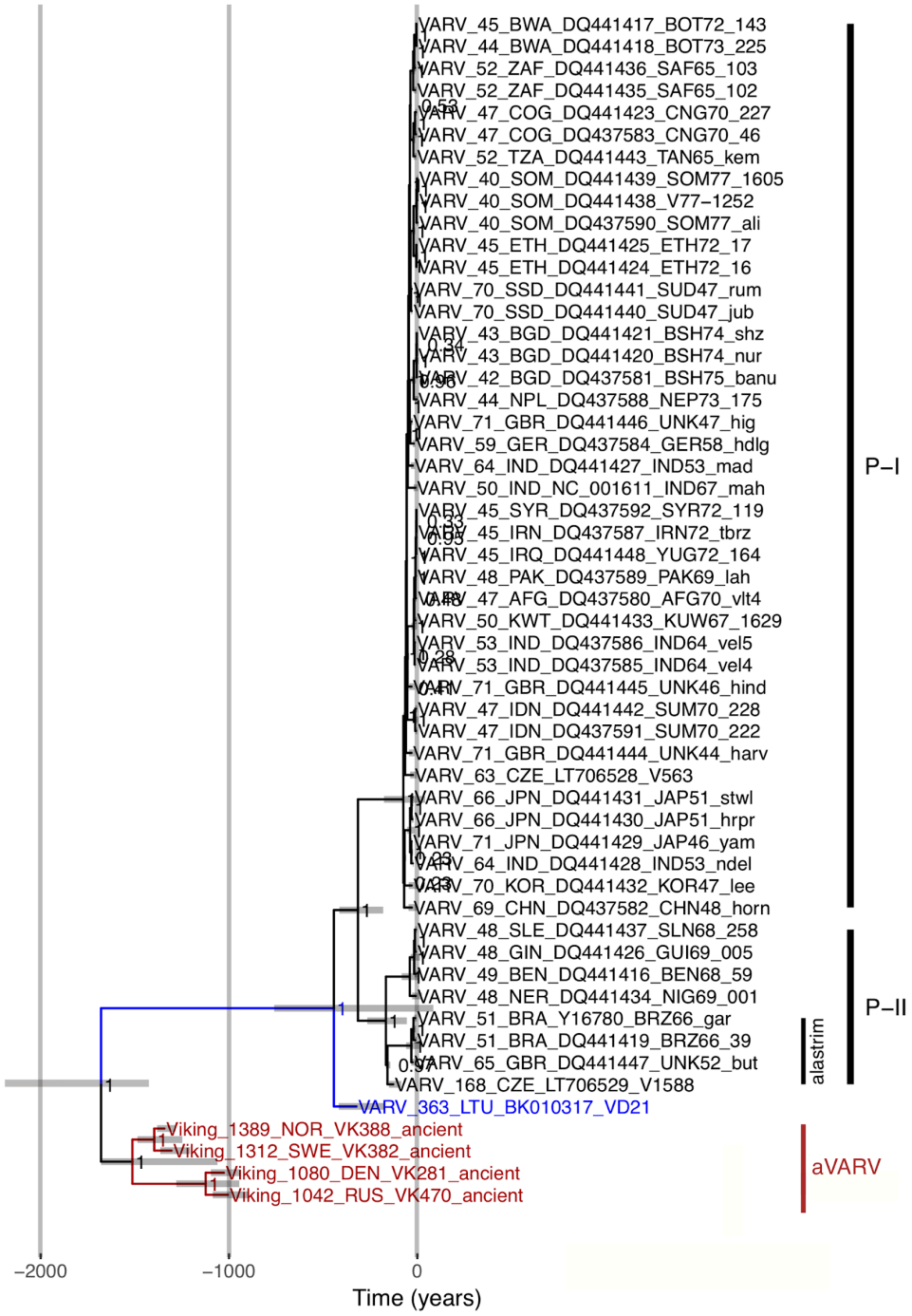


Fig. S15. Maximum clade credibility tree inferred with BEAST2. A log-normal relaxed clock and coalescent Bayesian skyline population prior are used, as well as a TPM1 substitution model with unequal base frequencies, invariant sites, and gamma distributed rate heterogeneity among sites. The substitution rate is inferred as 5.16×10^{-6} substitutions per site per year (s/s/y) (95% HPD interval: 3.67×10^{-6} to 6.51×10^{-6} s/s/y), and the root age as 1679 years old (yo) (95% HPD interval: 1423 to 2225 yo). The x-axis denotes time into the past, in years. The aVARV

sequences are shown in red, the ancient sequence from Duggan *et al.*, 2016 in blue (*16*). Taxon name fields indicate: virus/historical period, sample age relative to 2017, country abbreviation of sequence origin and region of sequence origin, as determined by the Standard country or area codes for statistical use, GenBank accession number/sample identifier, strain name. Horizontal grey bars indicate 95% HPD intervals of node ages. P-I and P-II refers to VARV primary clade I and primary clade II, respectively (*173*). 'alastrim' refers to alastrim minor, a phenotypically mild smallpox in America and West Africa (*173*).

A

B

C

D

E

Fig. S16A-E. Similarity, presence, and inactivation of genes in the ancient samples. A) VARV-VD21, B) aVARV-VK382, C) aVARV-VK388, D) aVARV-VK281, E) aVARV-VK470. Genes are shown on the x-axis. The figures are split in two parts. The bottom part shows the presence (filled circle), absence (empty circle) or inactivation (plus sign) of genes in black in the 17 orthopoxvirus reference sequences from Dataset 3. The top part shows the similarity of the reads from an ancient sample to the modern gene sequence. If an ancient sample had reads covering more than 50% of the reference gene sequence and the median bit score for the gene was within 2 of the highest median bit score for the gene, a coloured dot is shown in the row of the modern virus in question. The row labelled 'Highest' indicates the reference sequence with the highest bit score for each particular protein. A black square indicates that the modern reference sequence of a gene is inactive. A plus sign indicates a gene-inactivating mutation in the ancient consensus sequence, and a minus sign indicates an uncertain gene-inactivating mutation in the ancient consensus sequence. Black triangles in the top two rows indicate known host range (*11*) and diploid genes, respectively. Five genes on either side of the genome are greyed out and were not considered in other analyses, since they are diploid and identical. Gene labels correspond to the end offset of the gene in CPXV-Gri/GER, the name of the VACV-COP homolog, if applicable, separated by '|'.

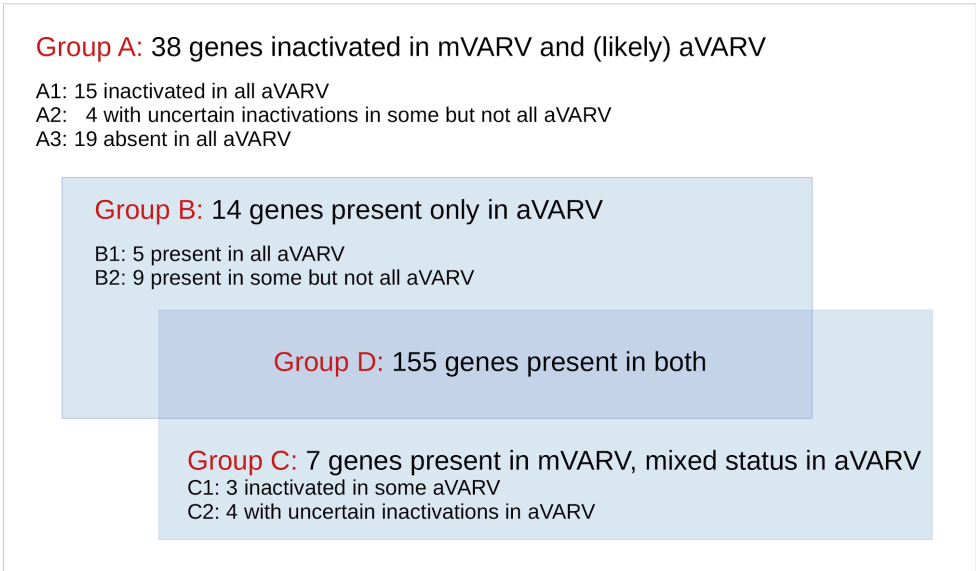


Fig. S17. Gene inactivation and presence groups. The figure shows a Venn diagram that divides 214 genes into four groups (A-D), according to their presence in mVARV and/or aVARV (blue rectangular regions) or absence in both (white background). In groups A and B, the status in the aVARV sequences is not always identical and in some cases is uncertain, so those groups are subdivided to provide additional detail. Details of inter- and intra-clade inactivation status for the 40 genes in subgroups A1 and A2 and groups B and C are shown in Fig. 2. Additional discussion of these individual genes, along with the names of all genes in all groups, are provided in the "Description of gene functions and mutations in the aVARV sequences" section above.

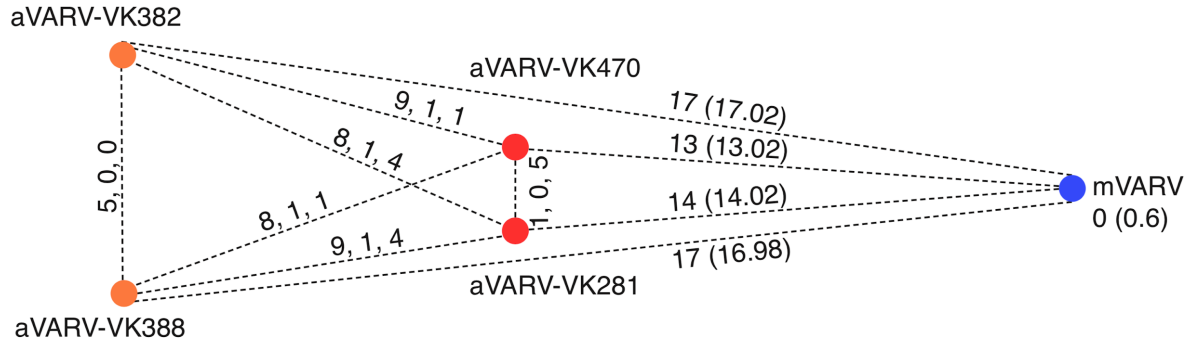


Fig. S18. Differences in gene activation among aVARVs and between aVARV and mVARV. Edges between the four aVARV viruses give the number of gene status differences for three categories of difference. A: *Stop/Absent* vs. *Present*, B: *UncertainFunction* vs. *Stop*, C: *UncertainStop* vs. *Present*. Thus the total difference between aVARV sequences is at least the first of these numbers and possibly as high as the sum of all three. Edges between mVARV and aVARV clade viruses give the median (mean) difference in number of active genes. The mVARV node label gives the mean (median) pairwise differences in gene activation status between a set of 50 modern VARV sequences. Viruses are approximately ordered by age from oldest to youngest from left to right. The names of the genes differing between aVARVs are as follows (given as A;B;C for the three categories above, with 'aVARV-' prefixes removed for conciseness): **VK388 vs. VK382**: *A35R*, *A40R*, *CrmB*, *C1L*, *E5R; none; none*. **VK470 vs. VK281**: *A35R*; none; *A55R*, *F10L*, *I8R*, *B20R*, *210863*. **VK470 vs. VK388**: *22682*, *C2L*, *K1L, F3L, 168145, A57R, C1L, E5R*; *A9L*; *A55R*. **VK470 vs. VK382**: *22682, C2L, K1L, F3L, A35R, A40R, 168145, A57R, CrmB*; *A9L*; *A55R*. **VK281 vs. VK388**: *22682, C2L, K1L, F3L, A35R, 168145, A57R, C1L, E5R*; *A9L*; *F10L, I8R, B20R, 201863*. **VK281 vs. VK382**: *22682, C2L, K1L, F3L, A40R, 168145, A57R, CrmB*; *A9L*; *F10L, I8R, B20R, 201863*.

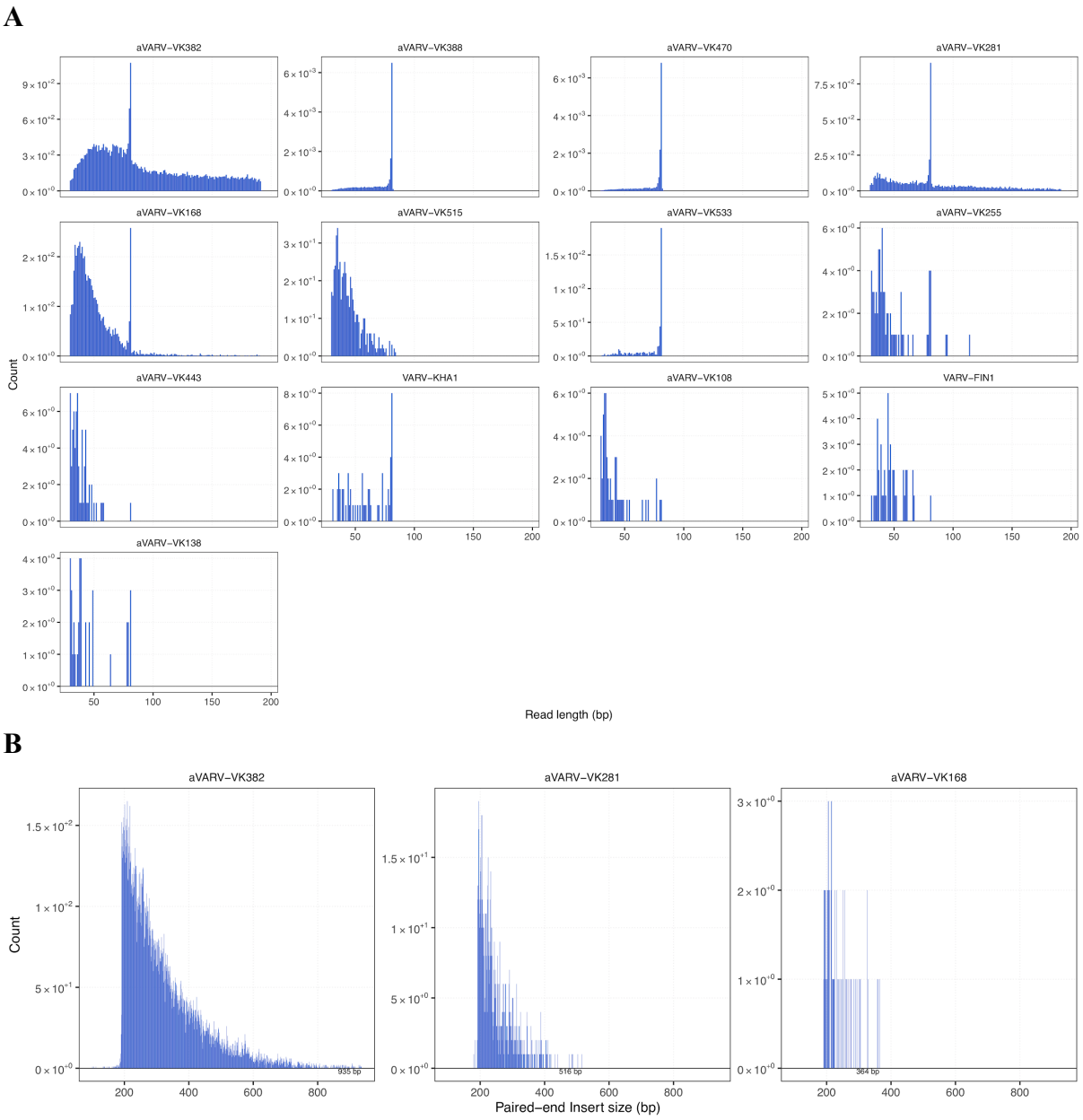
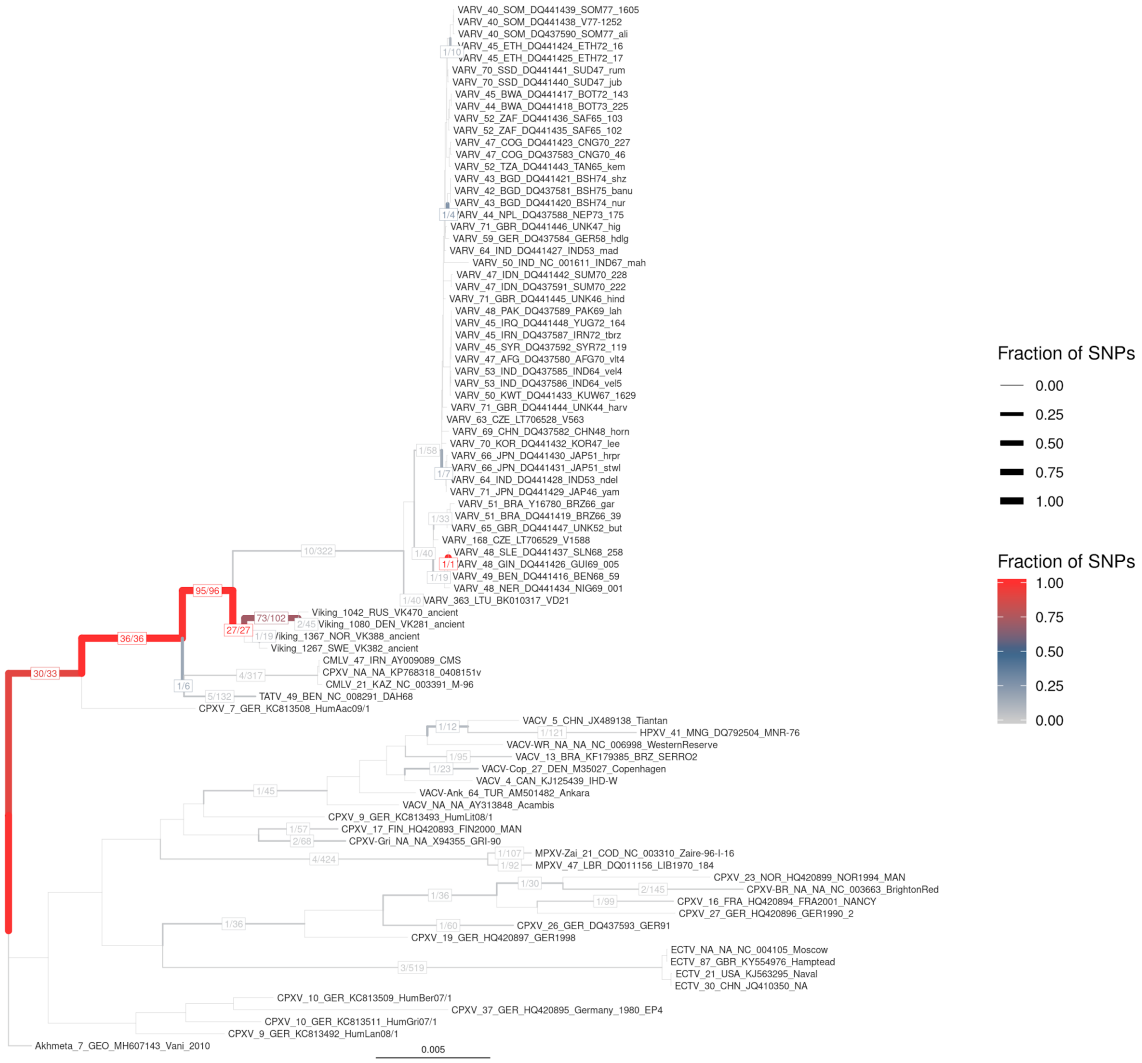


Fig. S19. Read length and insert size distributions of sequencing libraries. A) Read length distributions for single-end reads and collapsed paired-end reads. **B)** Insert size distribution of paired-end sequencing libraries, with maximum insert size labelled.

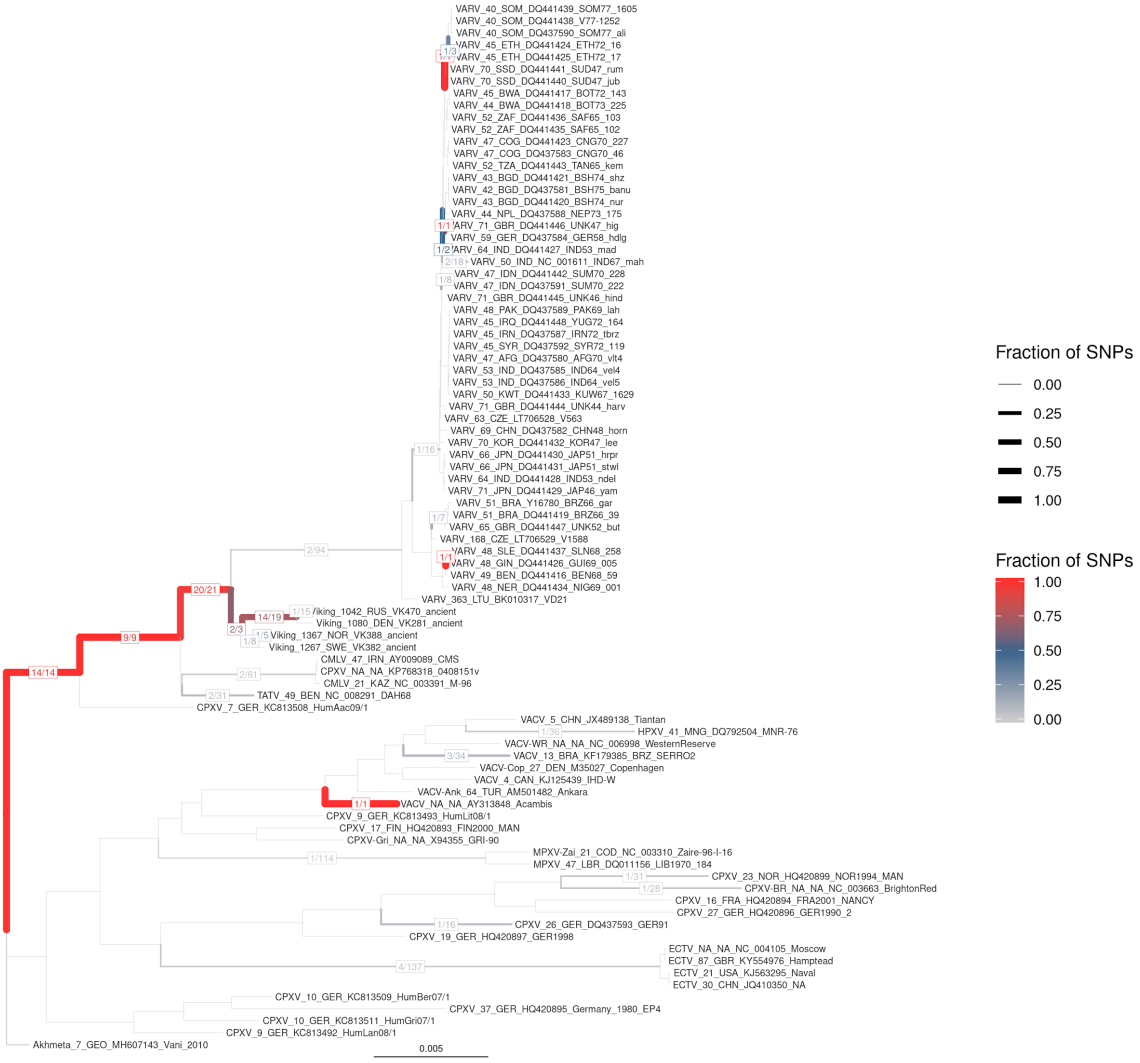
A

aVARV-VK168



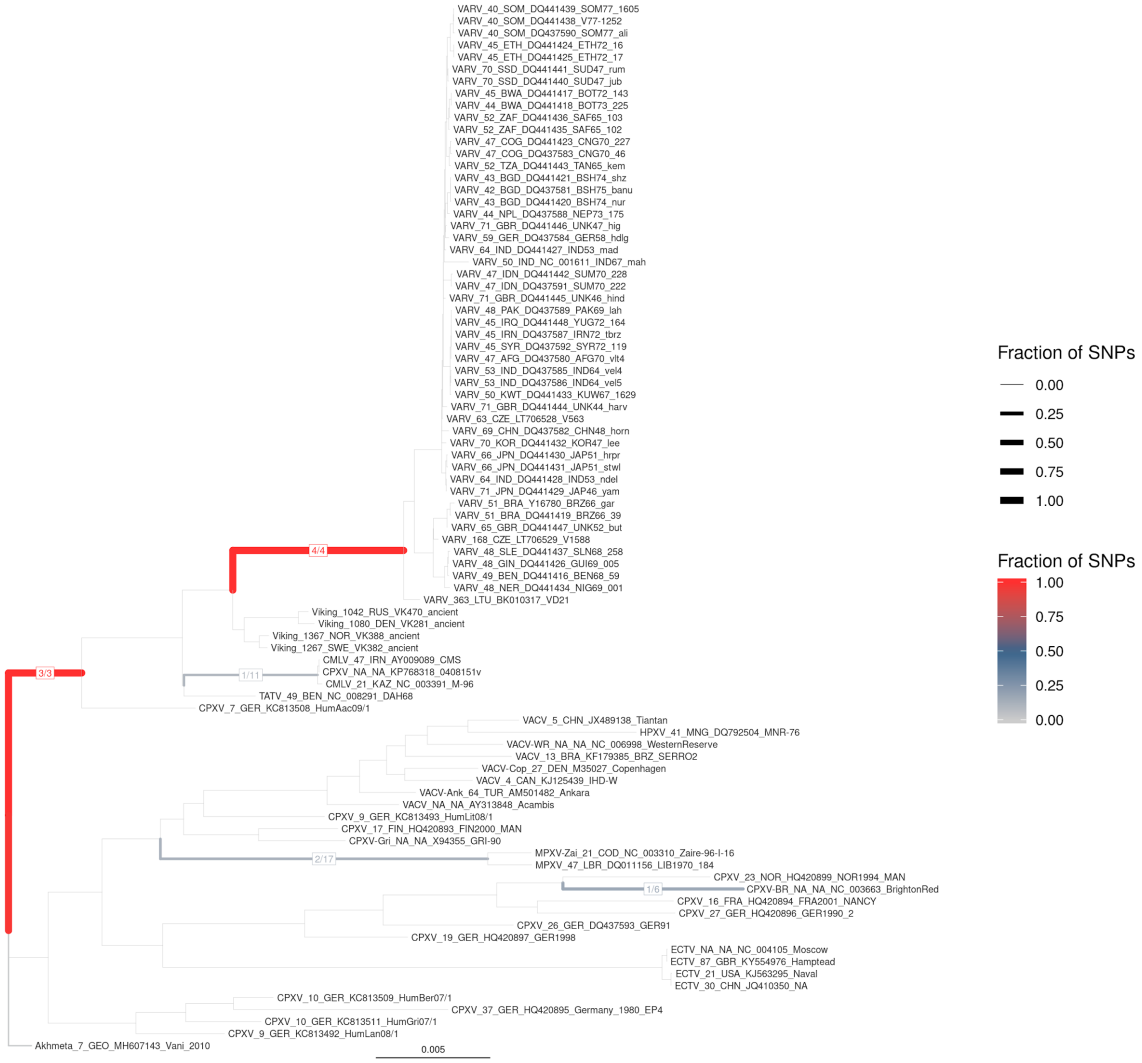
B

aVARV-VK533



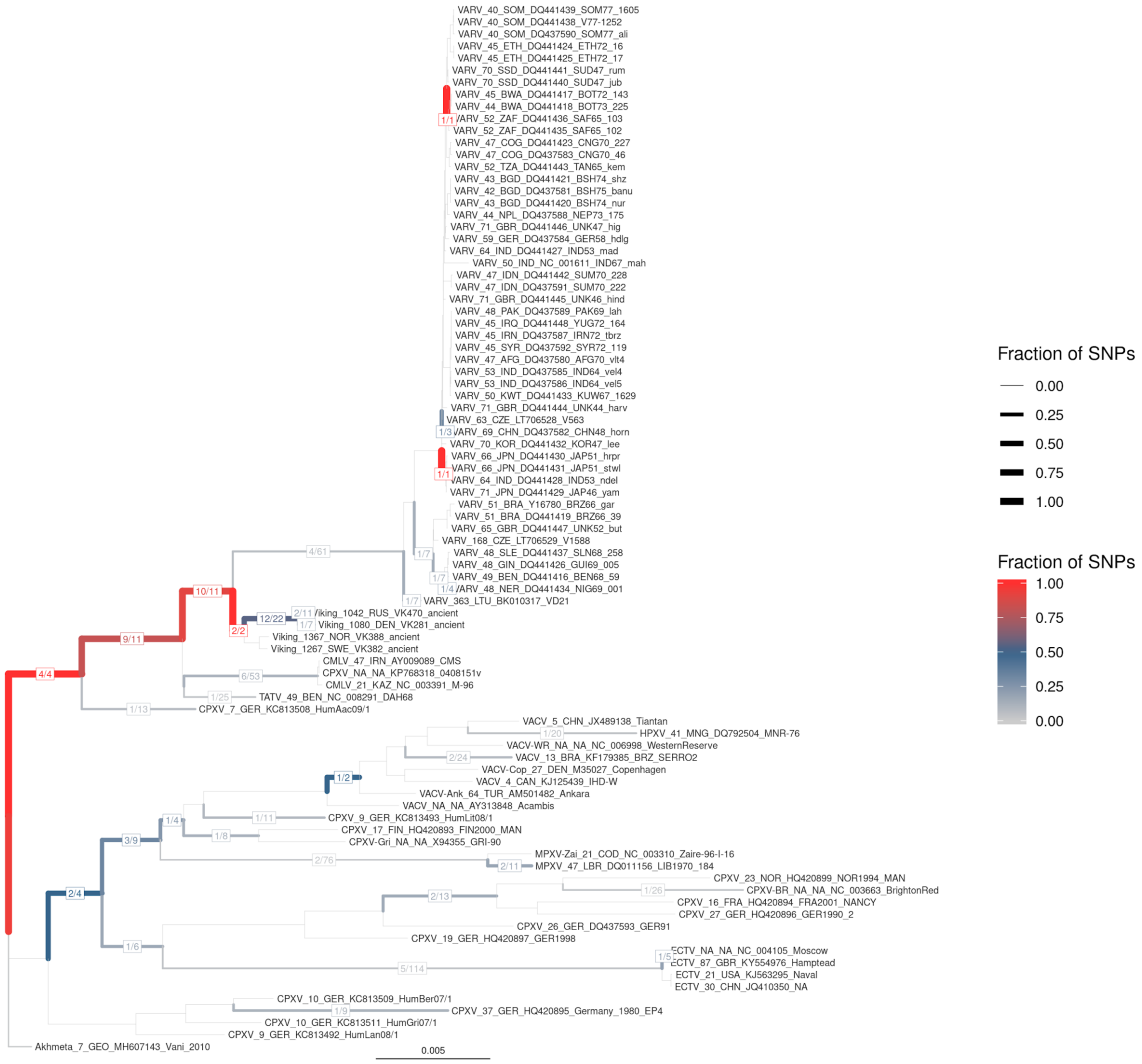
C

VARV-FIN1


VARV_40_SOM_DQ441439_SOM77_1605
VARV_40_SOM_DQ441438_V77-1252
VARV_40_SOM_DQ437590_SOM77_ali
VARV_45_ETH_DQ441424_ETH72_16
VARV_45_ETH_DQ441425_ETH72_17
VARV_70_SSD_DQ441441_SUD47_rum
VARV_70_SSD_DQ441440_SUD47_jub
VARV_45_BWA_DQ441417_BOT72_143
VARV_44_BWA_DQ441418_BOT73_225
VARV_52_ZAF_DQ441436_SAF65_103
VARV_52_ZAF_DQ441435_SAF65_102
VARV_47_COG_DQ441423_CNG70_227
VARV_47_COG_DQ437583_CNG70_46
VARV_52_TZA_DQ441443_TAN65_kem
VARV_43_BGD_DQ441421_BSH74_shz
VARV_42_BGD_DQ437581_BSH75_banu
VARV_43_BGD_DQ441420_BSH74_nur
VARV_44_NPL_DQ437588_NEP73_175
VARV_71_GBR_DQ441446_UNK47_hig
VARV_59_GER_DQ437584_GER58_hdlg
VARV_64_IND_DQ441427_IND53_mad
VARV_50_IND_NC_001611_IND67_mah
VARV_47_IDN_DQ441442_SUM70_228
VARV_47_IDN_DQ437591_SUM70_222
VARV_71_GBR_DQ441445_UNK46_hind
VARV_48_PAK_DQ437589_PAK69_lah
VARV_45_IRQ_DQ441448_YUG72_164
VARV_45_IRN_DQ437587_IRN72_tbrz
VARV_45_SYR_DQ437592_SYR72_119
VARV_47_AFG_DQ437580_AFG70_vlt4
VARV_53_IND_DQ437585_IND64_vel4
VARV_53_IND_DQ437586_IND64_vel5
VARV_50_KWT_DQ441433_KUW67_1629
VARV_71_GBR_DQ441444_UNK44_harv
VARV_63_CZE_LT706528_V563
VARV_69_CHN_DQ437582_CHN48_horn
VARV_70_KOR_DQ441432_KOR47_lee
VARV_66_JPN_DQ441430_JAP51_hrpr
VARV_66_JPN_DQ441431_JAP51_stwl
VARV_64_IND_DQ441428_IND53_ndel
VARV_71_JPN_DQ441429_JAP46_yam
VARV_51_BRA_Y16780_BRZ66_gar
VARV_51_BRA_DQ441419_BRZ66_39
VARV_65_GBR_DQ441447_UNK52_but
VARV_168_CZE_LT706529_V1588
VARV_48_SLE_DQ441437_SLN68_258
VARV_48_GIN_DQ441426_GUI69_005
VARV_49_BEN_DQ441416_BEN68_59
VARV_48_NER_DQ441434_NIG69_001
VARV_363_LTU_BK010317_VD21
4/4
Viking_1042_RUS_VK470_ancient
Viking_1080_DEN_VK281_ancient
Viking_1367_NOR_VK388_ancient
Viking_1267_SWE_VK382_ancient
CMLV_47_IRN_AY009089_CMS
3/3
1/11
CPXV_NA_NA_KP768318_0408151v
CMLV_21_KAZ_NC_003391_M-96
TATV_49_BEN_NC_008291_DAH68
CPXV_7_GER_KC813508_HumAac09/1
VACV_5_CHN_JX489138_Tiantan
HPXV_41_MNG_DQ792504_MNR-76
VACV-WR_NA_NA_NC_006998_WesternReserve
VACV_13_BRA_KF179385_BRZ_SERRO2
VACV-Cop_27_DEN_M35027_Copenhagen
VACV_4_CAN_KJ125439_IHD-W
VACV-Ank_64_TUR_AM501482_Ankara
VACV_NA_NA_AY313848_Acambis
CPXV_9_GER_KC813493_Hum08/1
CPXV_17_FIN_HQ420893_FIN2000_MAN
CPXV-Gri_NA_NA_X94355_GRI-90
2/17
MPXV-Zai_21_COD_NC_003310_Zaire-96-I-16
MPXV_47_LBR_DQ011156_LIB1970_184
CPXV_23_NOR_HQ420899_NOR1994_MAN
1/6
CPXV-BR_NA_NA_NC_003663_BrightonRed
CPXV_16_FRA_HQ420894_FRA2001_NANCY
CPXV_27_GER_HQ420896_GER1990_2
CPXV_26_GER_DQ437593_GER91
CPXV_19_GER_HQ420897_GER1998
ECTV_NA_NA_NC_004105_Moscow
ECTV_87_GBR_KY554976_Hamptead
ECTV_21_USA_KJ563295_Naval
ECTV_30_CHN_JQ410350_NA
CPXV_10_GER_KC813509_HumBer07/1
CPXV_37_GER_HQ420895_Germany_1980_EP4
CPXV_10_GER_KC813511_HumGri07/1
CPXV_9_GER_KC813492_HumLan08/1
Akhmeta_7_GEO_MH607143_Vani_2010
0.005
Fraction of SNPs
0.00
0.25
0.50
0.75
1.00
Fraction of SNPs
1.00
0.75
0.50
0.25
0.00

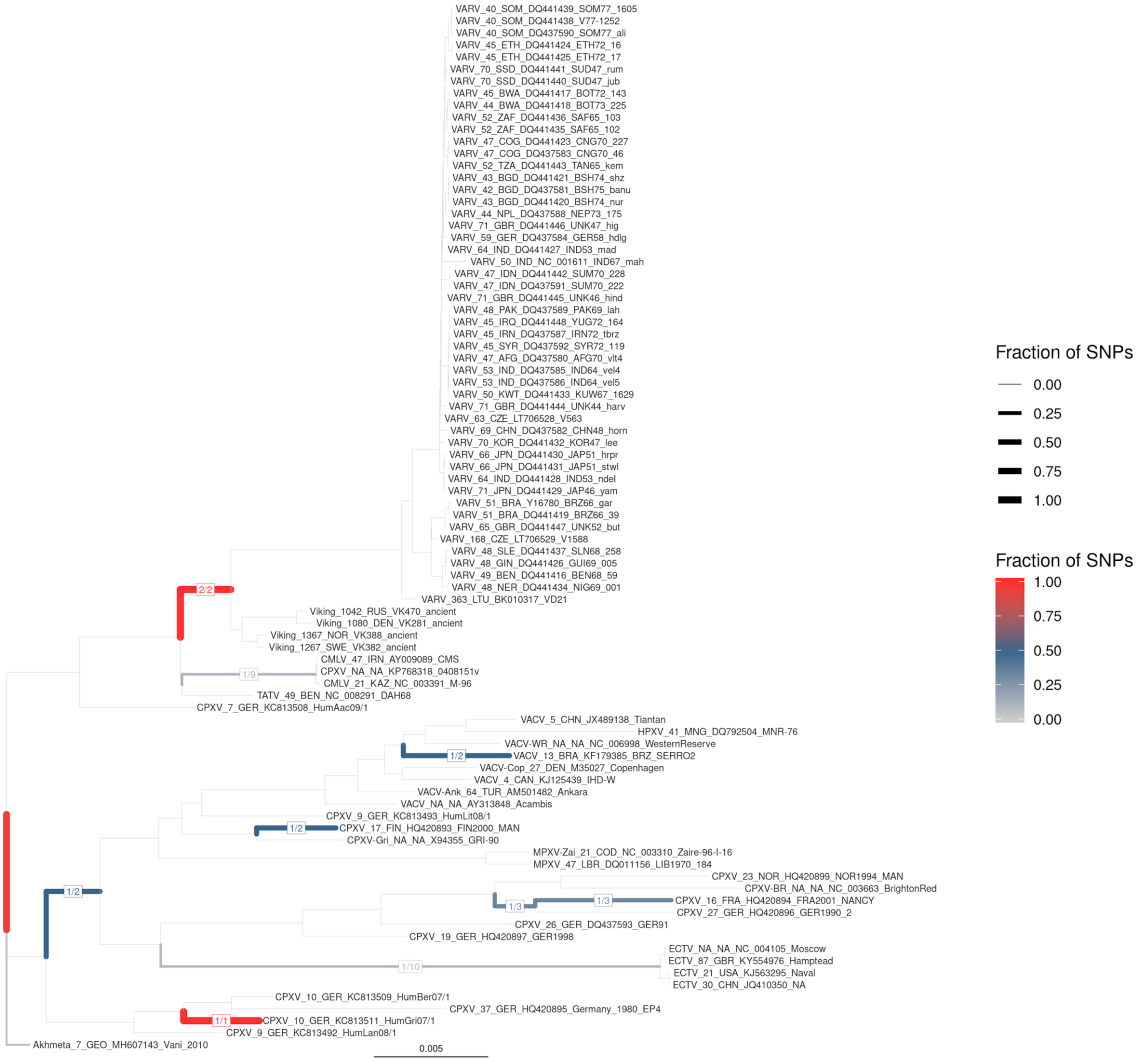
D

aVARV-VK515



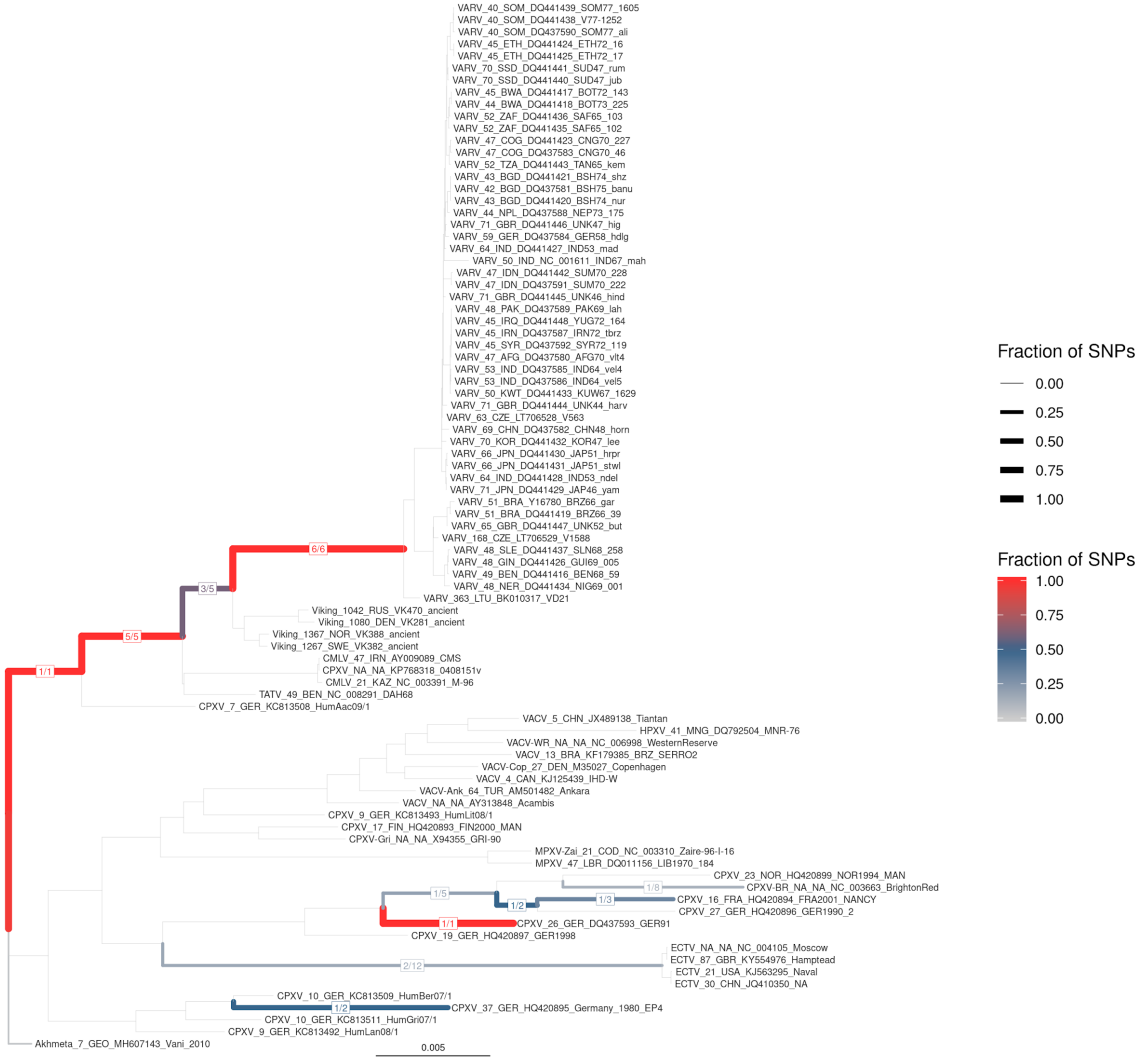
E

aVARV-VK443



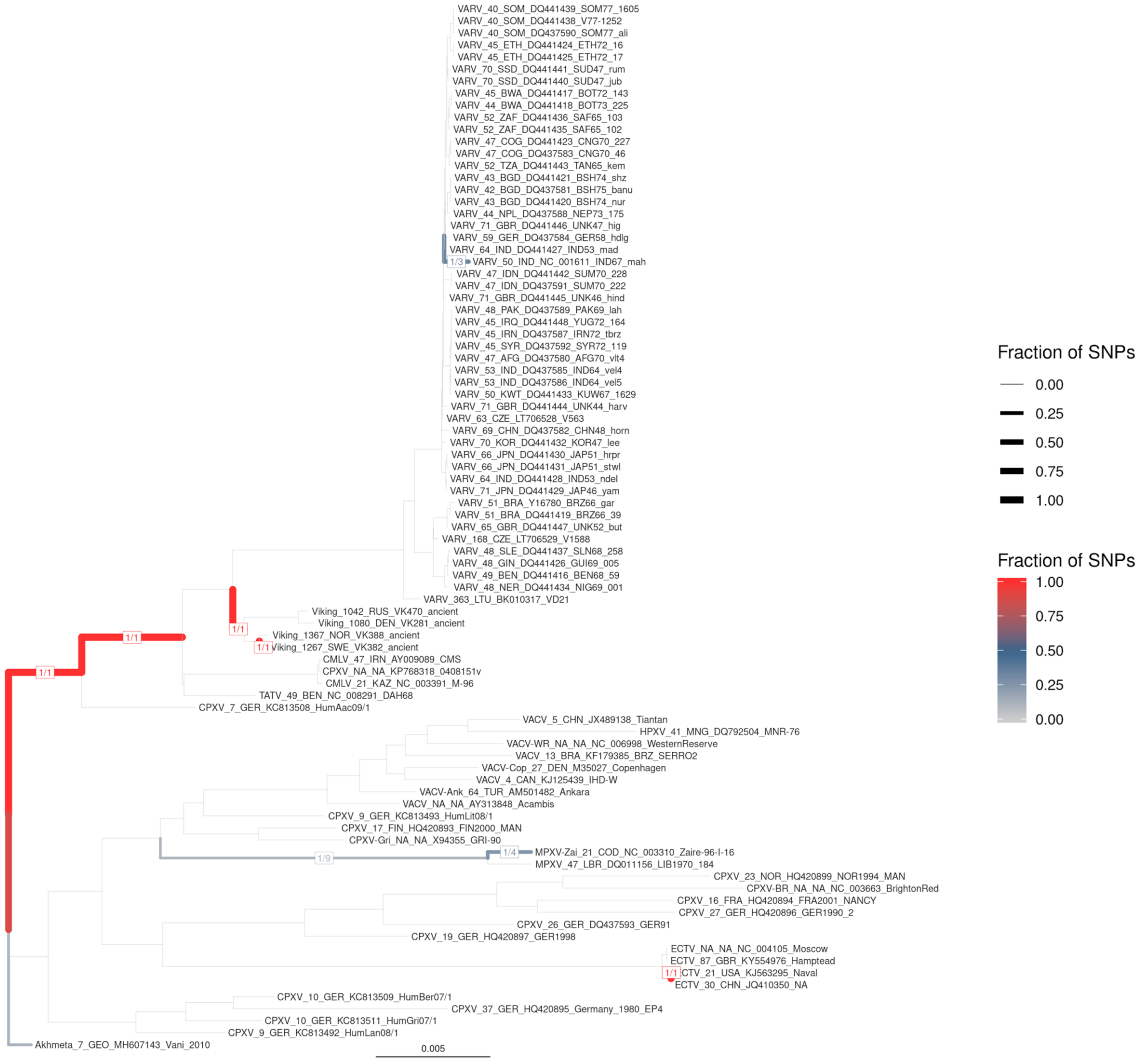
F

VARV-KHA1


VARV_40_SOM_DQ441439_SOM77_1605
VARV_40_SOM_DQ441438_V77-1252
VARV_40_SOM_DQ437590_SOM77_ali
VARV_45_ETH_DQ441424_ETH72_16
VARV_45_ETH_DQ441425_ETH72_17
VARV_70_SSD_DQ441441_SUD47_rum
VARV_70_SSD_DQ441440_SUD47_jub
VARV_45_BWA_DQ441417_BOT72_143
VARV_44_BWA_DQ441418_BOT73_225
VARV_52_ZAF_DQ441436_SAF65_103
VARV_52_ZAF_DQ441435_SAF65_102
VARV_47_COG_DQ441423_CNG70_227
VARV_47_COG_DQ437583_CNG70_46
VARV_52_TZA_DQ441443_TAN65_kem
VARV_43_BGD_DQ441421_BSH74_shz
VARV_42_BGD_DQ437581_BSH75_banu
VARV_43_BGD_DQ441420_BSH74_nur
VARV_44_NPL_DQ437588_NEP73_175
VARV_71_GBR_DQ441446_UNK47_hig
VARV_59_GER_DQ437584_GER58_hdlg
VARV_64_IND_DQ441427_IND53_mad
VARV_50_IND_NC_001611_IND67_mah
VARV_47_IDN_DQ441442_SUM70_228
VARV_47_IDN_DQ437591_SUM70_222
VARV_71_GBR_DQ441445_UNK46_hind
VARV_48_PAK_DQ437589_PAK69_lah
VARV_45_IRQ_DQ441448_YUG72_164
VARV_45_IRN_DQ437587_IRN72_tbrz
VARV_45_SYR_DQ437592_SYR72_119
VARV_47_AFG_DQ437580_AFG70_vlt4
VARV_53_IND_DQ437585_IND64_vel4
VARV_53_IND_DQ437586_IND64_vel5
VARV_50_KWT_DQ441433_KUW67_1629
VARV_71_GBR_DQ441444_UNK44_harv
VARV_63_CZE_LT706528_V563
VARV_69_CHN_DQ437582_CHN48_horn
VARV_70_KOR_DQ441432_KOR47_lee
VARV_66_JPN_DQ441430_JAP51_hrpr
VARV_66_JPN_DQ441431_JAP51_stwl
VARV_64_IND_DQ441428_IND53_ndel
VARV_71_JPN_DQ441429_JAP46_yam
VARV_51_BRA_Y16780_BRZ66_gar
VARV_51_BRA_DQ441419_BRZ66_39
VARV_65_GBR_DQ441447_UNK52_but
VARV_168_CZE_LT706529_V1588
VARV_48_SLE_DQ441437_SLN68_258
VARV_48_GIN_DQ441426_GUI69_005
VARV_49_BEN_DQ441416_BEN68_59
VARV_48_NER_DQ441434_NIG69_001
VARV_363_LTU_BK010317_VD21
Viking_1042_RUS_VK470_ancient
Viking_1080_DEN_VK281_ancient
Viking_1367_NOR_VK388_ancient
Viking_1267_SWE_VK382_ancient
CMLV_47_IRN_AY009089_CMS
CPXV_NA_NA_KP768318_0408151v
CMLV_21_KAZ_NC_003391_M-96
TATV_49_BEN_NC_008291_DAH68
CPXV_7_GER_KC813508_HumAac09/1
VACV_5_CHN_JX489138_Tiantan
HPXV_41_MNG_DQ792504_MNR-76
VACV-WR_NA_NA_NC_006998_WesternReserve
VACV_13_BRA_KF179385_BRZ_SERRO2
VACV-Cop_27_DEN_M35027_Copenhagen
VACV_4_CAN_KJ125439_IHD-W
VACV-Ank_64_TUR_AM501482_Ankara
VACV_NA_NA_AY313848_Acambis
CPXV_9_GER_KC813493_HumLit08/1
CPXV_17_FIN_HQ420893_FIN2000_MAN
CPXV-Gri_NA_NA_X94355_GRI-90
MPXV-Zai_21_COD_NC_003310_Zaire-96-I-16
MPXV_47_LBR_DQ011156_LIB1970_184
CPXV_23_NOR_HQ420899_NOR1994_MAN
CPXV-BR_NA_NA_NC_003663_BrightonRed
CPXV_16_FRA_HQ420894_FRA2001_NANCY
CPXV_27_GER_HQ420896_GER1990_2
CPXV_26_GER_DQ437593_GER91
CPXV_19_GER_HQ420897_GER1998
ECTV_NA_NA_NC_004105_Moscow
ECTV_87_GBR_KY554976_Hamptead
ECTV_21_USA_KJ563295_Naval
ECTV_30_CHN_JQ410350_NA
CPXV_10_GER_KC813509_HumBer07/1
CPXV_37_GER_HQ420895_Germany_1980_EP4
CPXV_10_GER_KC813511_HumGri07/1
CPXV_9_GER_KC813492_HumLan08/1
Akhmeta_7_GEO_MH607143_Vani_2010
6/6
3/5
5/5
1/1
1/8
1/5
1/3
1/2
1/1
2/12
1/2
0.005
Fraction of SNPs
0.00
0.25
0.50
0.75
1.00
Fraction of SNPs
1.00
0.75
0.50
0.25
0.00

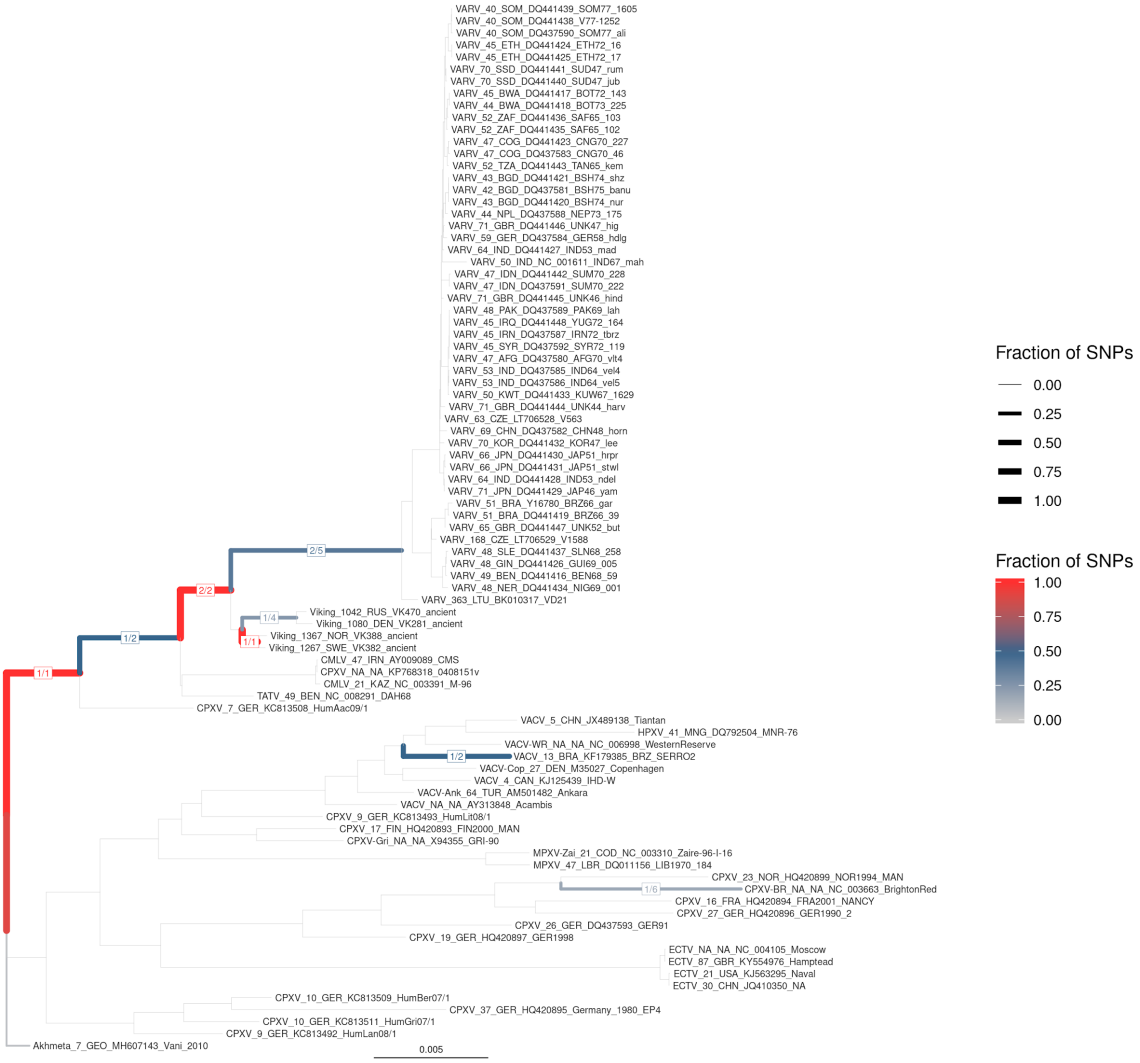
G

aVARV-VK108



H

aVARV-VK138



I

aVARV-VK255

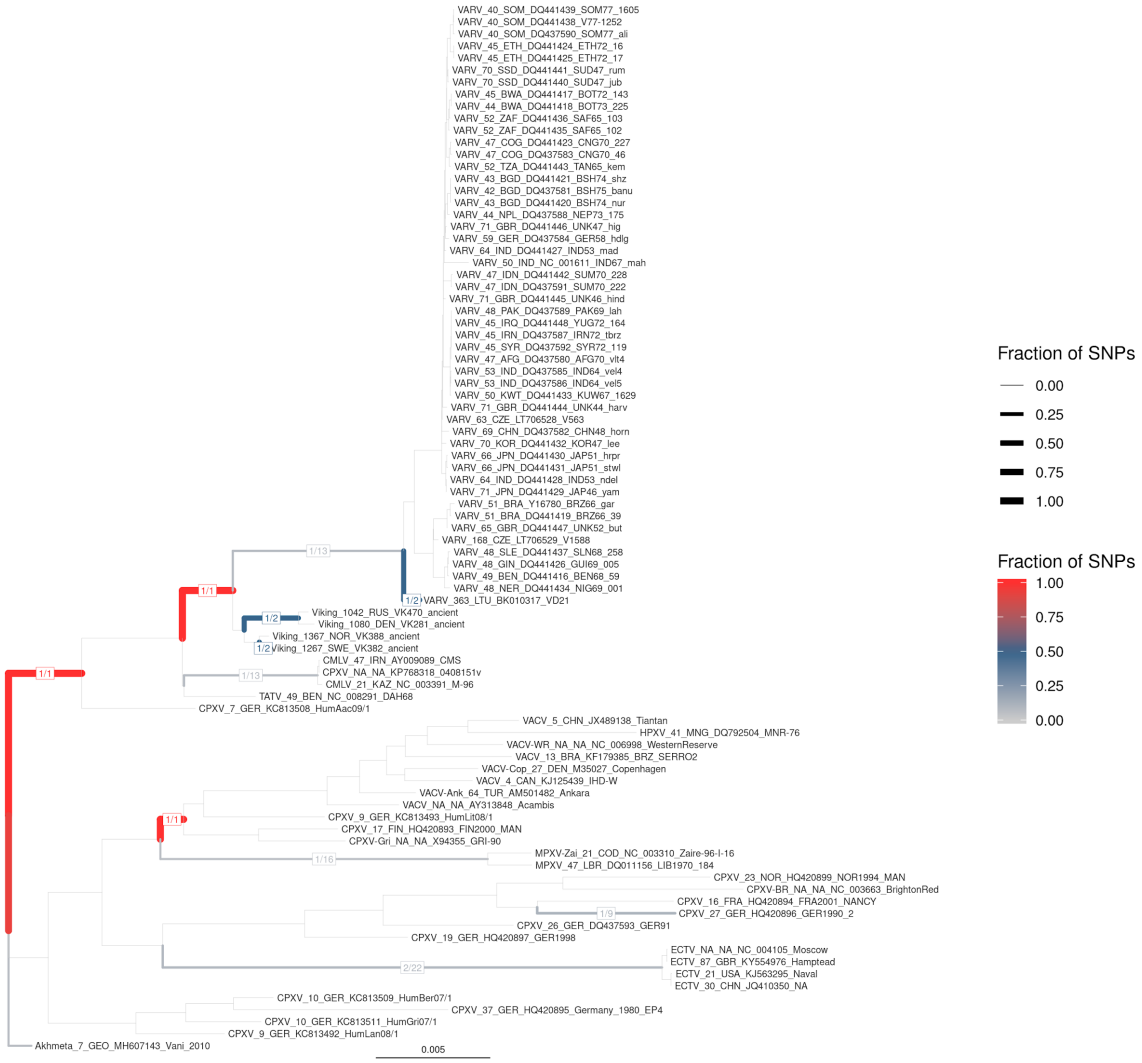


Fig. S20A-I. Branch-defining single-nucleotide polymorphism counts. A) aVARV-VK168, B) aVARV-VK533, C) VARV-FIN1, D) aVARV-VK515, E) aVARV-VK443, F) VARV-KHA1, G) aVARV-VK108, H) aVARV-VK138, I) aVARV-VK255. To complement and confirm the EPN-ng placements, branch-defining single nucleotide polymorphism (SNP) counts are indicated on branches as a pair of numbers, separated by a slash. A SNP is considered branch-defining if all sequences descending from the branch share the same nucleotide at that location, and all other sequences (on the alternate branches) do not. The first number gives the number of SNPs where the lower-coverage sample matches the branch-defining allele and the second gives the total number of SNPs where an allele is observed in the sample. For example, in panel A, the branch leading to the aVARV clade shows that out of the 27 branch-defining SNPs observed in the consensus aVARV-VK168, all 27 matched the allele defining that branch. Contrastingly, the branch leading to the VARV clade shows that that clade has 322 SNPs but the aVARV-VK281 consensus only matched the corresponding allele for ten of these. Thus, where

coverage exists in the ancient sequence, the differentiating SNP counts show agreement from the tree roots to the location where the EPA-ng algorithm places the ancient consensus. All trees are rooted with Akhmeta_7_GEO_MH607143_Vani_2010.

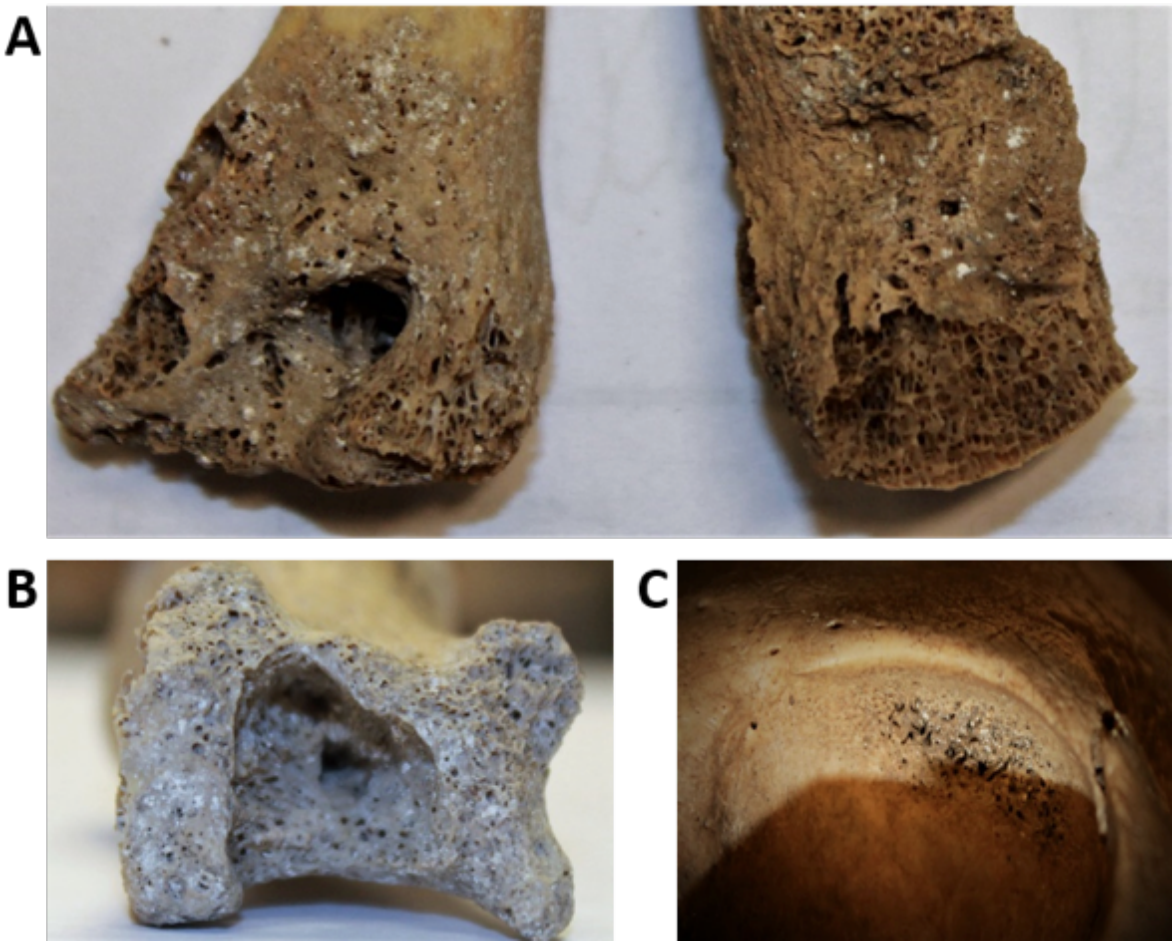


Fig. S21. VK388: Signs of pathology on the Nordland 253 specimen. A) Signs of infection on metatarsal. **B)** Bone resorption on metatarsal. **C)** Cranial cribra orbitalia observed in the orbital region.

Fig. S22. VK281: Horizontal artificially made groves (abrasion) on the labial side of the upper central incisors.

Fig. S23. VK443, grave 24 in SHM21367. Photo by T. J. Arne, 1935. ATA photo id7:181. Used by kind permission of ATA, digitized by Torbjörn Lennerud.

Fig. S24. VK533 excavation photo from 1966 of the burial A136. Note the small knife *in situ* on the clavicle. Courtesy of ATA (Antikvarisk Topografiska Arkivet), digitized by Torbjörn Lennerud and published with their kind permission.

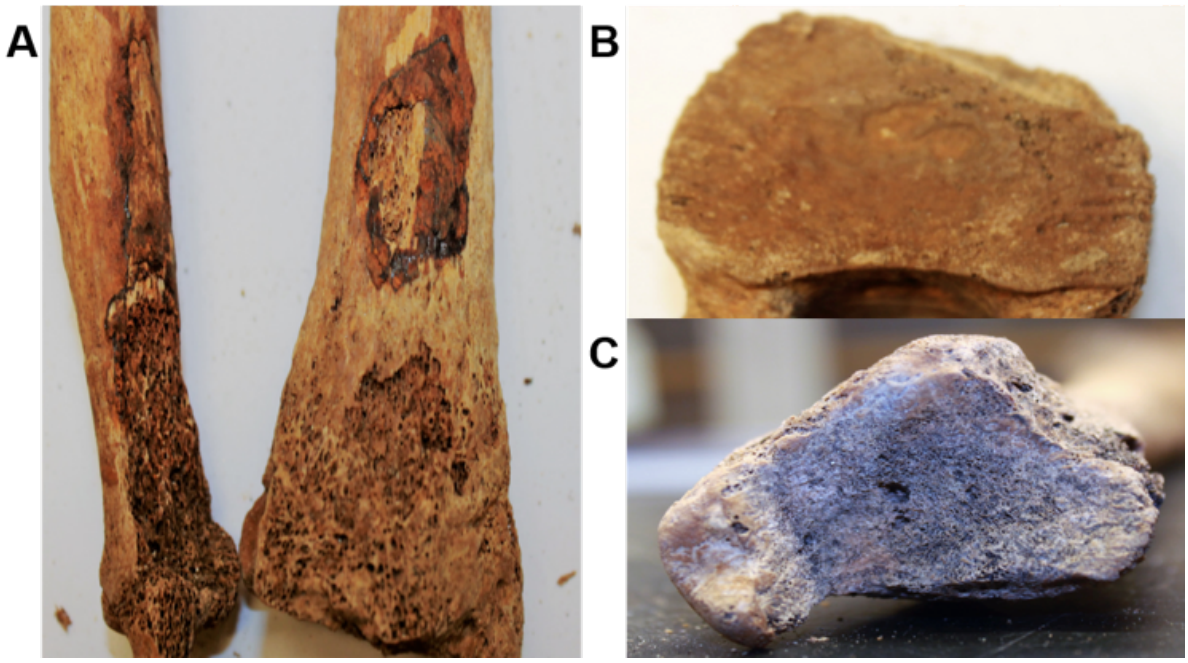


Fig. S25. Signs of pathology on the VK515 specimen. A) Signs of rust deposits and trauma on Ulna and Radius. **B)** Example of Schmorl's node on Lumbar vertebrae. **C)** Example of inflammatory pitting on tibia.

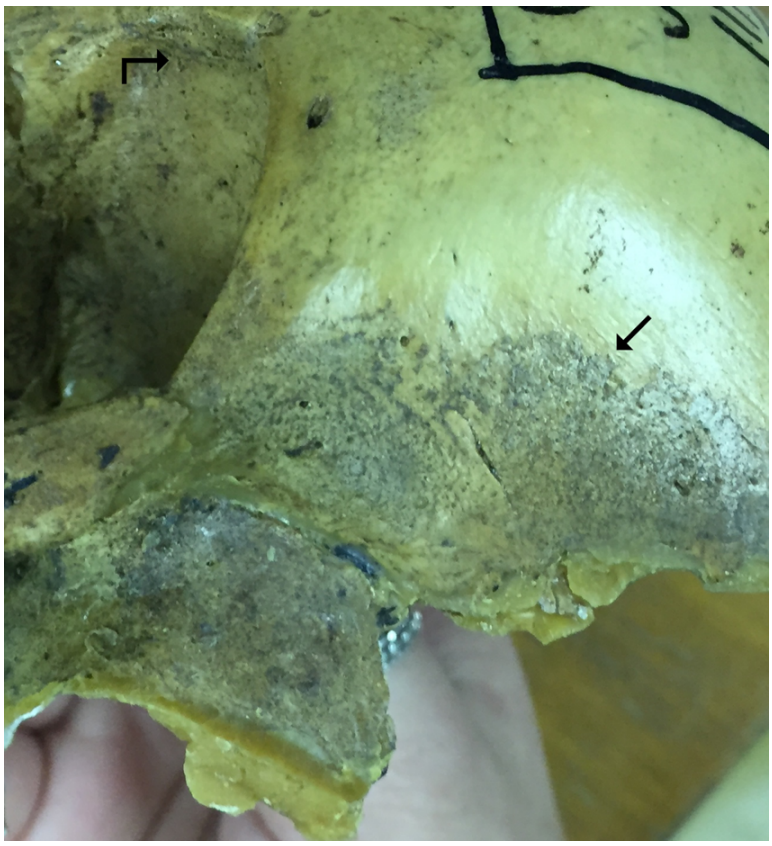


Fig. S26. KHA1: The upper left part of the calvarium of the child from Burial 66. Arrows identify porous hypertrophic bone formations on the upper border of the orbit and temporal area.

Fig. S27. The skull, frontal bone facing the camera, of 4A/PS/VK108. Note the thin crack in the bone above the left eye, passing through the brow ridge. This is most likely due to taphonomy and not a perimortem trauma as previously suggested (*122*). Photo: Helene Wilhelmson, Sydsvensk arkeologi AB.

Tables

Table S1. Mapping statistics. The spreadsheet contains three sheets with information on 1) per-library whole genome shotgun and capture sequencing statistics, 2) read coverage of reference sequences after mapping with Bowtie2, and 3) capture enrichment statistics.

Sample	Read count	Top hit orthopox	No 'nt' match	Top hit other
VK382	70,093	70,077	16	0
VK388	17,008	17,001	7	0
VK470	15,324	15,314	10	0
VK281	8537	8533	4	0
VK168	5834	5821	13	0
VK533	425	422	3	0
VK515	554	553	1	0
VK255	75	75	0	0
FIN1	45	45	0	0
KHA1	56	46	10	0
VK443	69	69	0	0
VK108	57	56	1	0
VK138	41	37	4	0

Table S2. Poxvirus specificity of reads. Reads matching TATV (GenBank accession no.: NC_008291.1) using Bowtie2 --end-to-end matching with (SAM) MAPQ >= 30 filtering were checked for specificity using BLASTn, by matching them against the NCBI 'nt' nucleotide database. The best-matching database sequence for each read was checked to see if it corresponded to an orthopoxvirus. The columns show: the sample name; the number of reads (shotgun plus capture, see Table S1 for breakdown) matching the TATV reference (these are lower than the TATV counts in Table 1 due to the use of --end-to-end instead of --local option in Bowtie2 and the MAPQ >= 30 filtering); the number of those reads that matched an orthopoxvirus as their top hit; the number of reads that did not match anything in the 'nt' database; the number of reads whose top match (if any) was not a orthopoxvirus. The table is sorted by mean depth of coverage (not shown), as in Table 1. See Materials and Methods for discussion.

Table S3. Gene annotations in ancient consensus. Table showing the gene name, minimum and maximum offset of the gene, gene length in nucleotides, and gene direction for each gene in the ancient consensus (*6*). Each worksheet in the table corresponds to one sample.

	Closest CMLV	Closest TATV	Closest modern VARV	VARV-VD21	aVARV-VK281	aVARV-VK382	aVARV-VK388	aVARV-VK470
VARV-VD21	0.985	0.987	0.997	1				
aVARV-VK281	0.978	0.981	0.978	0.979	1			
aVARV-VK382	0.99	0.993	0.989	0.99	0.986	1		
aVARV-VK388	0.989	0.992	0.988	0.989	0.985	0.998	1	
aVARV-VK470	0.987	0.989	0.986	0.987	0.987	0.994	0.994	1

Table S4. Sequence similarity between VARV-VD21, aVARV-VK281, aVARV-VK382, aVARV-VK388, and aVARV-VK470 consensus sequences and relevant orthopox reference sequences. Nucleotide similarity fraction was calculated for the central region (VACV-COP *F4L* to VACV-COP *A24R)* of the genome only.

Sequence	Suggested breakpoint offsets	Detecting algorithms (p-value)
aVARV-VK470	344 (130-385) and 468 (424-513)	3Seq ($7.7x10^{-3}$), BootScan ($2.4x10^{-14}$), GENECONV ($8.3x10^{-33}$)

Table S5. RDP4 recombination analysis. An RDP4 analysis found uncertain evidence of recombination in a 125-nucleotide region of the aVARV-VK470 sequence. The suggested breakpoint offsets are for the ungapped sequence and indicate the most likely breakpoint locations, with the 99% confidence interval of the breakpoints in parentheses. Detecting algorithms are followed by their average p-value in parentheses. In the possible recombination event, aVARV-VK281 was the suggested major parent with an unknown minor parent. The RDP4 program warned that "*the proposed recombination signal may be attributable to a process other than recombination*" and of a "*possible misidentification of recombinant*". No recombination signal was detected with aVARV-VK281, aVARV-VK382, or aVARV-VK388 as the recipient sequence. No modern sequence was identified as being the possible descendant of a recombination involving any ancient sequence. See Materials and Methods for additional details.

Model		Strict, coalescent constant	Strict, coalescent exponential	Strict, Bayesian skyline	Log-normal relaxed, coalescent constant	Log-normal relaxed, coalescent exponential	Log-normal relaxed, Bayesian skyline
	marginal L estimate	-153074.49	-153074.06	-153066.75	-153000.8	-153000.72	-152994.75
Strict, coalescent constant	-153074.49	0	0.43	7.74	73.69	73.77	79.74
Strict, coalescent exponential	-153074.06	-0.43	0	7.31	73.26	73.34	79.31
Strict, Bayesian skyline	-153066.75	-7.74	-7.31	0	65.95	66.03	72
Log-normal relaxed, coalescent constant	-153000.8	-73.69	-73.26	-65.95	0	0.08	6.05
Log-normal relaxed, coalescent exponential	-153000.72	-73.77	-73.34	-66.03	-0.08	0	5.97
Log-normal relaxed, Bayesian skyline	-152994.75	-79.74	-79.31	-72	-6.05	-5.97	0

Table S6. Model testing for different clock models and population priors. Models were compared using Path Sampling, as implemented in BEAST2. Likelihood values were compared using a Bayes factor test. A positive value for the Bayes factor implies support for the column model, a negative value support for the row model. According to Kass and Raftery, a Bayes factor in the range of 3–20 implies positive support, 20–150 strong support, and >150 overwhelming support (*81*). 'Strict' refers to a strict molecular clock model, 'Log-normal relaxed' to a log-normal relaxed molecular clock model. 'Coalescent constant' refers to a coalescent constant population prior, 'coalescent exponential' to a coalescent exponential population prior, and 'Bayesian skyline' to a coalescent Bayesian skyline population prior. 'Marginal L estimate' refers to the marginal log-likelihood estimate.

Model (molecular clock, population prior)	Root	aVARV clade	mVARV clade	VARV clade	Rate (s/s/y) $x10^{-6}$
Strict, coalescent constant	1702 (1586, 1836)	1534 (1491, 1585)	437 (405, 469)	335 (303, 370)	4.98 (4.42, 5.55)
Strict, coalescent exponential	1708 (1591, 1837)	1537 (1492, 1586)	437 (406, 471)	336 (305, 371)	5.97 (4.4, 5.53)
Strict, Bayesian skyline	1709 (1582, 1831)	1537 (1490, 1585)	438 (407, 472)	337 (307, 374)	4.96 (4.39, 5.52)
Log-normal relaxed, coalescent constant	1627 (1426, 1989)	1496 (1396, 1628)	438 (354, 577)	309 (212, 409)	5.36 (4.11, 6.55)
Log-normal relaxed, coalescent exponential	1756 (1447, 2345)	1536 (1408, 1742)	438 (350, 587)	308 (202, 419)	4.99 (3.53, 6.34)
Log-normal relaxed, Bayesian skyline	1679 (1423, 2225)	1513 (1400, 1689)	443 (349, 592)	315 (207, 432)	5.16 (3.67, 6.51)
Previously published					
Duggan *et al.*, 2016 (*16*), Strict, coalescent constant			372–429	224–283	
Smithson *et al.*, 2017 (*10*)			454–568	345–457	
Pajer *et al.*, 2017 (*17*)			667	322	
Porter *et al.*, 2017 (*18*)				372–463 / 375–419	

Table S7. Median root ages (in years into the past) and substitution rates inferred using BEAST2 under different clock models and population priors, compared to previously published estimates. Numbers in parentheses indicate 95% highest posterior density intervals. A TPM1 substitution model with unequal base frequencies, invariant sites, and gamma distributed rate heterogeneity among sites was used throughout. The substitution rate is given in substitutions per site per year (s/s/y). 'Strict' refers to a strict molecular clock model, 'Log-normal relaxed' to a log-normal relaxed molecular clock model. 'Coalescent constant' refers to a coalescent constant population prior, 'coalescent exponential' to a coalescent exponential population prior, and 'Bayesian skyline' to a coalescent Bayesian skyline population prior.

Table S8. Gene status and coverage in each sample. Worksheets show gene number as in Hendrickson et al., 2010 (*6*), VACV-COP name, gene status (*present*, *absent*, *noCoverage* (the gene does not have coverage), *UncertainFunction* (it is uncertain whether the gene is functional, since it is of intermediate length between a functional and a non-functional reference sequence), *UncertainStop* (uncertain if a gene-inactivating mutation is present, since the coverage of the gene-inactivating mutation is less than three reads)), the reference sequence used, coverage, and additional remarks. Each worksheet corresponds to one higher-coverage sample.

Site	Id number	Sex	Age	Skeletal element	Area	Paleopathological remarks
Nordland	Schreiner 253	Biological male	12–17-year-old	*Cranium*	Orbit	Cribra orbitalia
Nordland	Schreiner 253	Biological male	12–17-year-old	*Dentes*	Maxilla and mandible	Caries and calculus
Nordland	Schreiner 253	Biological male	12–17-year-old	*Ulna* dexter	Olecranon	Non-specific infection
Nordland	Schreiner 253	Biological male	12–17-year-old	*Ulna* sinister	Distal	Inflammation
Nordland	Schreiner 253	Biological male	12–17-year-old	*Radius* sinister	Medial	Inflammation
Nordland	Schreiner 253	Biological male	12–17-year-old	*Fibula* sinister	Proximal	Non-specific infection
Nordland	Schreiner 253	Biological male	12–17-year-old	*Tibia* sinister	Proximal and distal	Non-specific infection
Nordland	Schreiner 253	Biological male	12–17-year-old	*Tibia* dexter	Proximal and distal	Non-specific infection
Nordland	Schreiner 253	Biological male	12–17-year-old	*Calcaneus* sinister	Plantar	Non-specific infection
Nordland	Schreiner 253	Biological male	12–17-year-old	*Metatarsal*	Medial/intermedial	Non-specific infection and bone resorbtion

Table S9. Signs of pathology on the VK388 (Nordland 253) specimen. Skeletal remains were examined by ocular osteological analysis, identifying signs of inflammation, infection and bone resorbtion as well as dental pathology.

Id	Species	$^{87}Sr/^{86}Sr$	$\delta^{18}O$ (apatite)	$\delta^{13}C$ (apatite)	$\delta^{13}C$ (collagen)	$\delta^{15}N$ (collagen)	Assemblage	Age	Comments
1055	Human	0.7104	-6.1	-14.2	-20.1	13.3	A4	15-35	Teeth but also skeletal elements.
1077	Human	0.7269	-5.8	-14.5	-19.8	13.1	A8	15-35	Teeth but also skeletal elements DNA id 379.
1202	Dog	0.72147	-7.4	-12.4	-19.8	12.1	A4		M2 mandible sampled.

Table S10. VK382: Isotope results from Nabberör. Compiled from (*96–98*). The assemblage number indicates where in the boat the find originated and corresponds to the original excavation's documentation.

Assem-blage	Unique individual	MNI / area	Age (years)	Preservation	Comments
A4	1055	1	15–35	6 teeth (34, 36, 37, 44, 45, 46), also in a mandible.	Tooth 37 (root closed), over 15 years of age. Teeth 46 and 36, very little wear. Unlikely to be older than 35 years.
A8	1077	2	15–35	4 teeth (33, 35, 46, 47) but also skeletal elements (hands and feet R and L, arm and leg R). Duplicate of Mt II L (second individual).	47, newly erupted. 46 attrition identical to that in 46 in 1055. Very similar age to 1055?
A10	1150	1	18+	1 tooth (38) but also skeletal elements including the mandible. Skull fragments with perimortem sharp force trauma. Possible male, only skull fragments.	Third molar erupted, ie 18+ years old.
A7	1151	6	6	2 teeth (65,36).	65 (?), resorbtion of root initiated. 36 newly erupted. Could be from 2 separate individuals of the same age group (or one aged 4 and one aged 6y) but are assumed to be from just one person as no teeth are actual duplicates.
A7	1152	6	6–15	3 teeth (26, 36, 46)	26 and 46, newly erupted, root not complete ie < 15y)? 36, root almost complete (i.e 9-11y). Morphology and size of 36 and 46 very similar in 1151 and 1152. Possibly related? Siblings? Second smallest molars.
A7	1153	6	6–20	2 teeth (26, 46)	26 and 46, root at least 2/3 finished. 46, more attrition than fx 46 in 1152 and 1151. Smallest molars.

A7	1154	6	25–60	3 teeth (26, 46, 47) Also possibly (16, 36)	47, 26, 46 have complete roots and heavy attrition
A7	1155	6	9–30	-16 Also possibly (27)	16, very large, much larger than 26 in 1153 and 1156. Attrition slightly more than 16 in 1156, moderate so <30 y. Root complete (9+y). Largest teeth.
A7	1156	6	9–30	(16, 26) Also possibly (27, 37)	26, root complete (9+y). attrition moderate, <30y. Second largest teeth
A11	?	1	Adult	No teeth. Clavicle, Mc II L.	

Table S11. VK382: Summary of the human remains from the Nabberör burial. A7 is the centre of the boat, the location of the looting pit. The teeth in A7 were collected as one unit, with no observations made in the field of any separated units (matching a separation in different skulls or mandible and maxilla, for example). In A7 the division of teeth as separate individuals is based on preservation details (colour etc.), morphology, development and attrition (age), and size. They represent a low minimum of potential individuals in the assemblage A7. A further 34 teeth in A7 could not be assigned confidently to one single identified individual specified here. These teeth could match more than one of the individuals and therefore cannot be confidently attributed to just one of them. However, they may also come from even more individuals than it is possible to establish using MNI alone, and an osteological methodology (using ancient DNA, for example). FDI= definition of tooth by position in the dentition given by World Dental Federation (Fédération Dentaire Internationale) standard or ISO 3950. L=left, R=right.

DNA id	Id	Assemblage (anläggning)	Age (years)	Sex (osteology)	Sex (DNA)
VK379	Oland_1077	A8	15–35	Undetermined	XY
VK359	Oland_1130	A7	Over 15	Undetermined	XY
VK381	Oland_1131	A7	Over 15	Undetermined	XY
VK382	Oland_1132	A7	Over 15	Undetermined	XY

Table S12. VK382 aDNA results (*99*).

Epiphysis	Fusion	Corresponding age interval (male, years)	Source
Annual ring (lumbar vertebrae)	Partial	14–23	Cardoso & Rios 2011
Medial clavicle	Partial	16–30	Review in Cunningham et al. 2016: 261
Sacrum, S1-S2	Partial	17–-30	Cunningham et al. 2016
Iliac crest	Partial	15–24	Cunningham et al. 2016
Summary		**17–23**	

Table S13. A summary of the age estimation from epiphyseal fusion for VK443.

Table S14. VK533: Summary of all the East-West oriented burials from Öland. From Wilhelmson 2017a (*96*)**.** * ÖJG refers to the four volumes detailing Iron Age Graves in Öland:
Beskow-Sjöberg, M. H., & Arnell, K-H. (Eds.). (1987). Ölands järnåldersgravfält, Volym I. Stockholm: Riksantikvarieämbetet och Statens historiska museer.
Beskow- Sjöberg, M., & Hagberg UE. (Eds.) (1991). Ölands järnåldersgravfält, Volym II. Kalmar: Högskolan i Kalmar.
Hagberg, U. E., & Beskow-Sjöberg, M. (Eds.). (1996). Ölands järnåldersgravfält, Volym III. Stockholm: Riksantikvarieämbetet och Statens historiska museer.
Fallgren, J-H., & Rasch, M. (Eds). (2001). Ölands järnåldersgravfält, Volym IV. Stockholm: Riksantikvarieämbetet och Statens historiska museer.

BP refers to years before the present, relative to 1950.
** date as reported in excavation report (Hagberg 1966 (*109*)) and by Beskow-Sjöberg (1987) (*110*).

Grave id	Individual id	Date, as in ÖJG* (CE)	C14- Calibrated	C14 conventional, BP	Orien tation	Burial form
6	1038	0–200			NS	lime cist
108 F231 III	1014	0–200			NS	lime cist
108:I:169	1105	NA	853±71 CE	1175±50	NS	lime cist
108:IV:282	1106	NA	91±61 BCE	2065±45	NS	lime cist
134	1058	800–1050	1049±58 CE	1005±45	EW	pit
136	1076	150/-60-250/-60; period V:1			EW	pit
164	1075	800–1050	853±67 CE	1175±45	NS	lime cist

Table S15. VK533: A summary of the burials from Sörby-störlinge. Details available from Wilhelmson 2017a (*96*)**.** NS=North-South, EW=East-West. NA, not available. * ÖJG refers to the four volumes detailing Iron Age Graves in Öland:
Beskow-Sjöberg, M. H., & Arnell, K-H. (Eds.). (1987). Ölands järnåldersgravfält. Volym I. Stockholm: Riksantikvarieämbetet och Statens historiska museer.
Beskow- Sjöberg, M., & Hagberg UE. (Eds.) (1991). Ölands järnåldersgravfält, Vol. II. Kalmar: Högskolan i Kalmar.
Fallgren, J-H., & Rasch, M. (Eds). (2001). Ölands järnåldersgravfält.Vol. 4. Stockholm: Riksantikvarieämbetet och
Statens historiska museer.
Hagberg, U. E., & Beskow-Sjöberg, M. (Eds.). (1996). Ölands järnåldersgravfält. Volym III. Stockholm: Riksantikvarieämbetet och Statens historiska museer.

Site	Id no.	Biological sex	Age (years)	Skeletal element	Area	Paleopathological remarks
Nordland	Schreiner 4512	male	18–22	Dentes	Dentes	Enamel hypoplasia
Nordland	Schreiner 4512	male	18–22	Dentes	Dentes	Caries/Calculus
Nordland	Schreiner 4512	male	18–22	Radius sinister	Scaphoid/lunate facet	Inflammatory pitting
Nordland	Schreiner 4512	male	18–22	Lumbar vertebrae	Vertebral body	Schmorl's node
Nordland	Schreiner 4512	male	18–22	Tibia sinister	Inferior articular surface	Active inflammation
Nordland	Schreiner 4512	male	18–22	Tibia dexter	Inferior articular surface	Active inflammation
Nordland	Schreiner 4512	male	18–22	Fibula sinister	Malleolus	Inflammation distal
Nordland	Schreiner 4512	male	18–22	Fibula dexter	Malleolus	Inflammatory pitting
Nordland	Schreiner 4512	male	18–22	Lunate sinister	Superolateral aspect	Inflammatory pitting
Nordland	Schreiner 4512	male	18–22	Scaphoideum sinister	Convex surface	Inflammatory pitting

Table S16. Signs of pathology on the VK515 specimen. Skeletal remains were examined by ocular osteological analysis, identifying signs of dental pathology and inflammation.

Grave	Form	Orien-tation	Features	Artefacts	Com-ment	Position of bones
4/26 general Noted as mixed in the boxes of bones today present in the museum			See 4A, additional human remains in west part of the feature	no		
4A	Irregular, pit, 2.05 x 1.80 m	S-N	In the west part of the grave fill for A were the mixed remains of one individual (male, 169cm)	Iron: Key (17 cm); Key (?) (16.6 cm); Knife (15 cm); ring-shaped object (diameter 2.2 cm), object (length 4.3 cm). Burned bones (human), 1g. Unburned human bone.	Over Grave 4B.	Articulated skeleton.
4B, but noted 4C on the boxes of bones in the museum	Rectang-ular, 2.05 x 1.00 m	W-E	Coloration indicating wooden coffin under skeleton (1.65 x 0.65 m)	Iron: Knife (11.1 cm). Hairpin (?) (3.0 cm), Whetstone, slate (18.0 cm), Loom, soapstone (3.4 cm). Needle/pin? (bone) (4.1 cm), Unburned human bone.		Articulated skeleton. Found under 4A

Table S17. VK108: Description of grave 4/26. All measurements indicate length unless otherwise indicated. The text in the table is based on documentation in Malmö Museum, including notes from the archives.